

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK nucleic - nucleic search, using sw model

Run on: June 24, 2004, 07:32:31 ; Search time 10171 Seconds
(without alignments)
12680.637 Million cell updates/sec

Title: us-10-645-655-1

Perfect score: 4319
Sequence: 1 TCAATAGTCGTTTACTAGT.....ACCTTTTATTCATATAT 4319

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

EST:*
1: em_estba:*
2: em_estbun:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query length	DB ID	Description
C 1	89.2	2.1	1202	28	BH854408 303 Genom
C 2	58.8	1.4	1101	29	AL063921 Drosophi1
C 3	58.4	1.4	994	13	BX414650
C 4	56.6	1.3	885	13	BX425603

C 5	56.2	1.3	712	13	BX416727	BX416727
C 6	55.6	1.2	1124	13	BX36282	BX36282
C 7	53.4	1.2	900	28	AZ533096	AZ533096
C 8	53.2	1.2	999	23	BX380865	BX380865
C 9	52.4	1.2	968	29	CNS06PBX	CNS06PBX
C 10	52.4	1.2	1201	13	BX461310	BX461310
C 11	51.2	1.2	926	14	CB986430	CB986430
C 12	50.8	1.2	617	14	CA743799	CA743799
C 13	50.2	1.2	770	28	BZ231524	BZ231524
C 14	50	1.2	1101	29	CNS017Y	CNS017Y
C 15	50	1.2	1111	13	BX376338	BX376338
C 16	49.6	1.1	872	28	AZ530255	AZ530255
C 17	49.6	1.1	1201	13	BX376097	BX376097
C 18	49.4	1.1	869	13	BX401089	BX401089
C 19	49.4	1.1	1164	13	BX425839	BX425839
C 20	49.2	1.1	1055	13	B0876453	B0876453
C 21	49	1.1	753	13	BX443342	BX443342
C 22	49	1.1	919	29	CNS005R2	CNS005R2
C 23	48.8	1.1	475	12	BM029212	BM029212
C 24	48.8	1.1	977	13	BX397883	BX397883
C 25	48.6	1.1	892	28	BZ256924	BZ256924
C 26	48.4	1.1	711	29	CNS011V8	CNS011V8
C 27	48.4	1.1	828	29	CNS011TX	CNS011TX
C 28	48.4	1.1	1201	29	CNS016E1	CNS016E1
C 29	48.4	1.1	1306	13	B0675354	B0675354
C 30	48.2	1.1	633	13	BX403519	BX403519
C 31	48	1.1	1180	13	BX436369	BX436369
C 32	47.8	1.1	710	12	BM163311	BM163311
C 33	47.8	1.1	988	14	CD249057	CD249057
C 34	47.8	1.1	1005	14	CB207771	CB207771
C 35	47.8	1.1	1201	13	BX356851	BX356851
C 36	47.8	1.1	1201	13	BX46717	BX46717
C 37	47.6	1.1	545	29	CE364861	CE364861
C 38	47.6	1.1	558	29	CE133657	CE133657
C 39	47.6	1.1	728	29	CE697734	CE697734
C 40	47.6	1.1	1016	29	CNS016C9	CNS016C9
C 41	47.4	1.1	849	29	CNS03664	CNS03664
C 42	47.2	1.1	1099	13	BX456575	BX456575
C 43	47.2	1.1	1101	29	CNS0182P	CNS0182P
C 44	47	1.1	1101	29	CNS000FMC	CNS000FMC
C 45	46.8	1.1	1079	13	BX398967	BX398967

ALIGNMENTS

RESULT 1
BH854408/c 1202 bp DNA linear GSS 25-JUN-2002
LOCUS 303 Genomic fragments from E. coli EC317 absent from B. coli K-12
DEFINITION MG1655 Escherichia coli genomic 5', genomic survey sequence.

ACCESSION BH854408
VERSION BH854408.1 GI:21554464
KEYWORDS GSS.

ORGANISM

Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1 to 1202)
Stocki, S.L., Babluk, L.A., Rawlyk, N.L., Porter, A.A. and Allan, B.C.
Identification of Genomic Differences between E. coli strains
Pathogenic for Poultry and E. coli K-12 MG1655 using Suppression
Subtractive Hybridization Analysis
Unpublished (2002)

JOURNAL

COMMENT

Contact: Stacy Stocki
Veterinary Infectious Disease Organization
University of Saskatchewan
120 Veterinary Road, Saskatoon, Saskatchewan, Canada, S7N 5E3
Tel: 3069667467
Fax: 3069667478
Email: stocki@duke.usask.ca
Fragment 303
PCR Primers

REFERENCE 1 (bases 1 to 1101)

1080 AAGATAAAGCTTCTAATCTAGATATGACGGACCTAATATTTATCTCCACCGTTAAC 1139

Db 502 TWAAWTAATAAAAAAATTTTTTTTTTTWAAWTAATAATTTTTTTTAAAT 443

Oy	1140	AATGGAAGACGCATTGTTCCTTGGATCAAAAACAAGCATTAATCTTGCAATC	1195
Dd	442	TTTTTTTWTMTAAATTTATTTTTCCTAAAAAAMMMWMTATATKCCCCCCTC	387
 RESULT 3 BX144650/c			
LOCUS	BX144650	994 bp	mRNA linear EST 15-MAY-2003
DEFINITION	BX144650 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAF001YN02		
VERSION	BX144650		3-PRIME, mRNA sequence.
KEYWORDS	BX144650.1 GI:30763455		EST.
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Ll,M.B., Gruber,C., Tessee,J. and Polayes,D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001) Contact: Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seque@genoscope.cns.fr Web : www.genoscope.cns.fr Library was constructed by life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6015.f Contact : Peng liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSOCAF001DG01NP1. Location/Qualifiers 1..994 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSOCAF001YN02" /tissue_type="THYMUS" /clone_lib="Homo sapiens THYMUS" /note="vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. library was not normalized."		
 ORIGIN			
Query Match	1.4%; Score 58.4; DB 13; Length 994;		
Best Local Similarity	21.2%; Pred. No. 0.18;		
Matches 136;	Conservative 214; Mismatches 291; Indels 2; Gaps 1;		
Oy	824	AAATCGTAAATTAAGGGATTAATGCGGAAGCCAACCTTTGAAGCAAAGAAATG	883
Dd	968	WDDDSRKKDMMWAADAKMBAARDMMWAGAARFRPADMDMAAFMRWDMSTWWAAR	929
Oy	884	GTTTCATTTGTTGGCAATCTTATTTGATGCAATTTTCGAAGAAGATTACATACATC	943
Dd	928	KKAADTTTKTKKQDAADMGDDMDMTATMWAADMMALRAATATWTDMDTAW	869
Oy	944	ACTTACACCAGCTGTGATGAGGTGACACAAATTGAGGAATGATTAATGCTAGGG	1003
Dd	868	TRDXKATWADDADDDDKTKT--TTTWTMMKMTDVGWTAATWADATWADADAAR	811
Oy	1004	GTCATAACTGCAAACTGCGAATACCACTGAGAAATTAATTAATTCGTAAGCAATNGAG	1063
Dd	810	AMWAATAAMWRAPAAAMWAADADMDMMWMTMDAADAAAMWAAMDYTKRAMPADWAPAAA	751
Oy	1064	TTTACCTTTGAAAGAGAGAGATTAATGCCATATCCATGATGACGAGCACTAATATT	1123
Dd	750	AMWDWGRGBRADAGDKARAAWWDGAGRGDXDRBMAARMDDADABAARPAAMAAAATAW	691
Oy	1124	TTCTCCAGCTTTAACAAATGAGAAAGCCTATATTTTATGATCAAAAACAAGATATT	1185
Dd	690	ATWAKWKCDWMAAAADATKAGRKDADDDWMAAMTTTTTTTTTAAADAKMKMKWT	631

[illegible]

```

QY 2744 TTATTTGGCTATTAAGAGATTAATTAATTAATCAATGACGTGAGCGATACGT 2803
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 686 KXKXGKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 627
QY 2804 ATTATCTGTTGCAACACGAGCAAGAAACCCGAAACCTTGACAACTTAAGTTGTTGA 2863
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 626 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 567
QY 2864 AAGCAAGAATATCAACGTTATCAAGTAAAGCTCAATTTACTTTAGAAATGACACAGT 2923
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 566 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 507
QY 2924 TGATGACGTGCTATTAAGTTATTAATAGTGAAGAAATGATGCGAATTCGCTTGATTA 2983
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 506 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 447
QY 2984 CCCAATTAAGAGAGAGAAATTCACATATGATTTAGTAAAGCAGACAGACAGACGAA 3043
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 446 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 387
QY 3044 ATTAGAAGCCAAACAAGTTGAACGACTGCTAATAACAACAAGAGTGAAGCAAAAGTGG 3103
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 386 KTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 327
QY 3104 GTCAAGAAAGAGCAGCAGCAGCGCTTCTGATACCTGCTGATCAAGCCTGTTAA 3163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 326 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 267
QY 3164 CGCATTAAGAGCCAAACAAGCTGACTGACTGCTGAAACAAAAAAAGTAAAGCAAAAC 3223
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 266 TNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 207
QY 3224 AAAAAAAGTGGCTCAAAAAA 3245
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 206 AAAAAAAGTGGCTCAAAAAA 185

```

RESULT 5
 LOCUS BX416727/c 712 bp mRNA linear EST 15-MAY-2003
 DEFINITION BX416727 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
 CS0DA011114 5-PRIME, mRNA sequence.
 ACCESSION BX416727
 VERSION BX416727.1 GI:30765629
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 712)
 L1, W.B., Gruber, C., Jesse, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by life technologies, a division of
 Invitrogen. Contact: Feng Liang Email: fliang@life.techn.com URL:
 http://fulllength.invitrogen.com/InvitrogenCorporation1600
 Faraday Avenue Genoscope sequence ID: CS0DA011B070P1.
 Location/Qualifiers
 1..712
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DA011114"
 /tissue_type="NEUROBLASTOMA"
 /clone_lib="Homo sapiens NEUROBLASTOMA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo (GT) primer. Five prime end enriched
 double-strand cDNA was digested with NotI and cloned into
 the NotI and EcoRV sites of the pCMVSPORT 6 vector."

```

ORIGIN Library was not normalized.
Query Match 13%, Score 56.2, DB 13; Length 712;
Best Local Similarity 15.8%, Pred. No. 0.51;
Matches 71; Conservative 177; Mismatches 201; Indels 0; Gaps 0;

QY 3030 CAAGCAGAAAGCACTTGAAGCCAAACAAGTTGACGACTCTAAAAACAAACAGT 3089
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 712 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 653
QY 3090 GAGCCAAAAGTGCAGTGAAGAGAGCAGAGAGAGGTTCTGATACCTGCTGAT 3149
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 632 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 593
QY 3150 CAAAGCTGTTAAAGCATTAGAGCAACAAGTGAAGTGAAGTGAAGCAACAAAA 3209
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 592 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 533
QY 3210 AGTAAAGCAAAAACAAGGAGGCGGTCAAAAAAGACAGCTTTTCTGATCCCTGCTT 3269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 532 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 473
QY 3270 GATCAAGCCTGTTGCTATGAAGCCGCACTGAGTTAGTTGATGCCCAACAGATCG 3329
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 472 CAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 413
QY 3330 GAAAAAGATGCTTGAAGTCAAGAGAGGAAAAACAAGCAACAAAGAAAGCTGATC 3389
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 412 AAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 353
QY 3390 AGCGTTATTCAGATGAGTATGATGATGATGATGATGATGATGATGATGATGAT 3449
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 352 AAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 293
QY 3450 GTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3478
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 292 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 264

```

RESULT 6
 LOCUS BX436282 1124 bp mRNA linear EST 15-MAY-2003
 DEFINITION BX436282 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP001YCO1
 5-PRIME, mRNA sequence.
 ACCESSION BX436282
 VERSION BX436282.1 GI:30787521
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1124)
 L1, W.B., Gruber, C., Jesse, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by life technologies, a division of
 Invitrogen. Contact: Feng Liang Email: fliang@life.techn.com URL:
 http://fulllength.invitrogen.com/InvitrogenCorporation1600
 Faraday Avenue Genoscope sequence ID: CS0CAP001A01QPI.
 Location/Qualifiers
 1..1124
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0CAP001YCO1"
 /tissue_type="THYMUS"
 /clone_lib="Homo sapiens THYMUS"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed

with a NotI-oligo(NT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 1.3%; Score 55.6; DB 13; Length 1124;
Best Local Similarity 23.0%; Pred. No. 0.65;
Matches 172; Conservative 209; Mismatches 366; Indels 0; Gaps 0;

QY 1744 TTGCTACAGCGTGGTGGGAGACGATTAATAATATACCAATGGCGGATTAAACC 1803
DB 1017 YYYTAAAT 958
QY 1804 TAT 1863
DB 957 AAAAAAT 898
QY 1864 AAGGCAAT 1923
DB 897 AAAAAAT 838
QY 1924 CTTACAT 1983
DB 837 YYYTAT 778
QY 1984 TGTGGATCATCAT 2043
DB 777 YYYTAT 718
QY 2044 GAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2103
DB 717 YAAAAAT 658
QY 2104 ATGCAAT 2163
DB 657 AAAAAAT 598
QY 2164 ATTGACAT 2223
DB 597 YYYTAT 538
QY 2224 CTATACAT 2283
DB 537 YYYTAT 478
QY 2284 TTAAGCTTGAAT 2343
DB 477 YYYTAT 418
QY 2344 TAAACAT 2403
DB 417 AHTTAT 358
QY 2404 TGAAT 2463
DB 357 AAAAAAT 298
QY 2464 AAAAAAT 2490
DB 297 AAAAAAT 271

RESULT 7
A2533096/c 900 bp DNA linear GSS 03-NOV-2000
LOCUS ENT0D134TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, genomic survey sequence.
ACCESSION A2533096
VERSION A2533096.1 GI:11088690
KEYWORDS GSS.
SOURCE Entamoeba histolytica
ORGANISM Entamoeba histolytica

REFERENCE
AUTHORS Loftus, B., Van Aken, S. and Fraser, C.
TITLE Determination of clone end sequences from Entamoeba histolytica
JOURNAL Unpublished (2000)
COMMENT HML:IMSS sheared DNA library
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@igir.org

FEATURES
source
1. .900
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HML:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHS01; Site 1; Bet 1; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, C.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999)."

ORIGIN

Query Match 1.2%; Score 53.4; DB 28; Length 900;
Best Local Similarity 44.7%; Pred. No. 1.9;
Matches 250; Conservative 0; Mismatches 306; Indels 3; Gaps 1;

QY 2986 CATTAAGAGAGCAATATGACATGATTAGTAGAGAGAGAGAGAGAGAGACT 3045
DB 614 CAGCAG 555
QY 3046 TGAAGCCAAACAGTGAACCGACTCTAAACACAAACAGGTAGCCAAAGTCCGT 3105
DB 554 CTGAAG 495
QY 3106 CAG 3165
DB 494 CAAAG 435
QY 3166 CATTAAG 3225
DB 434 CAGCAG 375
QY 3226 AAAAAAT 3282
DB 374 AAG 315
QY 3283 TGGCATTAAG 3342
DB 314 AAGCAG 255
QY 3343 TAGCTAAG 3402
DB 254 AAGCAG 195

```

Qy 3403 ATATGCGGTATATGAGATTATCTGCAACAGTAATATGATGCTTCTTCTTCAAGATGAT 3462
Db 194 AAGATGACAAAGGCGAAGAAACCAACACCAAAATATACAGATATATGACCAACCA 135
Qy 3463 TGAATGCTCTTTTGTAGATCAAGCAATCTCCCTGTGACAAATATGACAGAGTA 3522
Db 134 AAGAAAGAGGAAAGAAATATCAACAAACCAAGAGATGAAATATAGTAACGACACCA 75
Qy 3523 AAAGACGCTATGATCTGA 3541
Db 74 AAGCAACAGTACTCTGA 56

RESULT 3
BX380865 999 bp mRNA linear EST 08-MAY-2003
LOCUS BX380865
DEFINITION BX380865 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
Clone CSD1058YE10 3-PRIME, mRNA sequence.
ACCESSION BX380865 GI:30439622
VERSION BX380865
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 999)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91036 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3221.r. For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSD1058BC05NP1&cluster=3221.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradey Avenue Genoscope sequence ID : CSD1058BC05NP1.
FEATURES
Source
location/Qualifiers
1..999
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSD1058YE10"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pOWS-PORT 6 vector. Library was normalized."
ORIGIN
Query Match 1.2%; Score 53.2; DB 13; Length 999;
Best Local Similarity 24.6%; Pred. No. 2;
Matches 112; Conservative 118; Mismatches 225; Indels 0; Gaps 0;
Qy 695 TGGAGCAGTAATGATATCGTATTTGGAGGCGATGTTGTTAAAGCGGAGATATG 754
Db 399 TGAASANSSTTANNSTANSANSSSSGSSCNSSSTSTNNSGANNSSAVAAATSS 458
Qy 755 TGCATTCGATTCGAGGCTCAAGAGGGGACATGTTCTCCGATGTTATTTATGATGC 814
Db 459 TSSAATNNSAASNTTBBTAAABAAATATBTNTTANNTTANNTBBBAADMTATTA 518
Qy 815 TGAAGAAACAAATGTTAATTAATGAGCATTAACGAGAGGACCCCTTTGAAGGCA 874
Db 519 TAAATKKAADAAANAGRAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 578
Qy 875 AGAAATGSGTTTCAATGTTTCCGAACTTATTTTATGAAATTTTGGAAAGATTT 934

```

```

Db 579 WWAATTTTWTWTATAKNDKKAAXAADADADADADADADADADADADADADADADADADADAD 638
Qy 935 ACATACATCACTTACACCCGACCTGTATGAGCTATACATTAATGATGATTA 994
Db 639 AADADAMADAKKKDAGDAGDAGDAGDAGDAGDAGDAGDAGDAGDAGDAGDAGDAG 698
Qy 995 TGGTCAGGCTCTATTAATCTGAGAAATACATGAGAAATTAATTAATTAATGATGAC 1054
Db 699 ADADADADADADADADADADADADADADADADADADADADADADADADADADAD 758
Qy 1055 AATATGAGTTTACCTTGAAGAGAGATTAAGTTTATCTATGATGAGAGCC 1114
Db 759 ADADADADADADADADADADADADADADADADADADADADADADADADADADAD 818
Qy 1115 TATATTTTCTTCCAGTTTAAACATGAGAAA 1145
Db 819 AKADADADDTKMAAANAKMAAAMATTTAAANAAA 853

RESULT 9
CNS06PBX 968 bp DNA linear GSS 05-JUL-2001
LOCUS CNS06PBX
DEFINITION T7 end of clone AV00A013B12 of library AV00A from strain CBS 379 of
Saccharomyces exiguus, genomic survey sequence.
ACCESSION AL409171 GI:12176428
VERSION AL409171
KEYWORDS GSS.
SOURCE Saccharomyces exiguus
ORGANISM Saccharomyces exiguus
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 968)
AUTHORS Souciet, J.L., Aigle, M., Artiguenave, F., Blandin, G.,
Bollotin-Pukhara, M., Bon, B., Brottier, P., Casaregola, S.,
de-Montigny, J., Dujon, B., Durans, P., Lepingle, A., Llorente, B.,
Malpertuy, A., Neuvéglise, C., Ozier-Kalogeropoulos, O., Potier, S.,
Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Jouvel, M.,
Wincker, P. and Weissenbach, J.
TITLE Genomic exploration of the hemiascomycetous Yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL FEMS Lett. 487 (1), 3-12 (2000)
Db 1152876
MEDLINE 20584711
PUBMED 1152876
2 (bases 1 to 968)
REFERENCE 2
AUTHORS Bon, E., Neuvéglise, C., Lepingle, A., Wincker, P., Artiguenave, F.,
Galliard, C. and Casaregola, S.
TITLE Genomic exploration of the hemiascomycetous yeasts: 6.
Saccharomyces exiguus
JOURNAL FEMS Lett. 487 (1), 42-46 (2000)
Db 20584716
MEDLINE 1152881
PUBMED 1152881
3 (bases 1 to 968)
REFERENCE 3
AUTHORS Genoscope.
TITLE Direct Submision
JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5736, 91057 EVRY cedex, FRANCE. (E-mail :
segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
FEATURES
Source
location/Qualifiers
1..968
/organism="Saccharomyces exiguus"
/mol_type="genomic DNA"
/strain="CBS 379"
/db_xref="taxon:34358"

```

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="CS0DF025YA16"
```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINT at: <http://image.llnl.gov>
 Plate: NDCM141 row: 1 column: 16

High quality sequence stop: 142.
FEATURES

Source
Location/Qualifiers
1..926
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="TMAG:30329463"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1ib="NIH MGCC 184"
/note="Organ: Pooled-glandular; Vector: pDR-LIB; Site_1: SfiI (ggccatagcc); Site_2: SfiI (ggccgcctggcc); Library is oligo-dT primed and directionally cloned. cDNA was prepared from a glandular pool of tissues from thyroid, parathyroid, adrenal, cortex and pineal gland. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATATGCGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGCGGCGCATG-dT(30)BN-3' (where S = A, C, or G and N = A, C, G, or T). Average insert size 1.38 kb (range 0.60-3.5 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech laboratories (Palo Alto, CA). Note: this is a NIH MGCC library."

ORIGIN

Query Match 1.2%; Score 51.2; DB 14; Length 926;
Best Local Similarity 44.1%; Pred. No. 5.2;
Matches 215; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

2758 AACGCGTAATTAATTAATTCATATGCGTGAAGGCGATTAATATATCTGTGCA 2817
|||||
559 AA 500
2818 ACACAGGCAAGACCCGAAACCTTGAGCAATTAATTTGTTGAAGCAAGATATC 2877
499 ACTAAGCAACAACAAACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 440
2878 AACGTTATGATGACTCAATTTACTTTGAAATGACACGTTGATGAGGTGCAT 2937
439 AAGAGAGAAACAAAGACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 380
2938 TACGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2997
379 AA 320
2998 AGAATTCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3057
319 AA 260
3058 AAGTTGACCGACTCTTAACACAAACGCTGACCAAGAGTCCGCTCAAGAGACG 3117
259 GAAA 200
3118 CGAGAGGAGGCTTCCGTATCCTGCTGATCAAGCGCTTTAAAGCATTAAGACCA 3177
139 AAAAAAAAAAAAAAAAAAGATCGAATTAAGAAATGAAAGAAAGAAAGAAAGAA 140
3178 AACAGCTGAATGACTGCTGAAACACAAAGAAAGTGAAGCAAAAGAAAGTGGCT 3237
139 AA 80
3238 CAAAAAGA 3245
Db 79 AAAAAA 72

RESULT 12
CA743799 617 bp rRNA linear EST 26-NOV-2002
LOCUS
DEFINITION wtls.pk005.p2 wtls Triticum aestivum cDNA clone wtls.pk005.p2 5'
ACCESSION CA743799
VERSION CA743799.1 GI:25559622
KEYWORDS EST.

SOURCE
ORGANISM
Triticum aestivum (bread wheat)
Triticum aestivum

REFERENCE
AUTHORS
1 (bases 1 to 617)
Tingey, S.V., Powell, W., Walters, P., Dolan, M., Hatney, C., Yuan, Z.,
Mao, G., Caraher, R. and Hanafey, M.K.

TITLE
JOURNAL
Unpublished (2002)
Contact: Scott V. Tingey
Crop Genetics
E. I. Dupont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@usa.dupont.com
Seq primer: 77.

FEATURES
source

Location/Qualifiers
1..617
/organism="Triticum aestivum"
/mol_type="rRNA"
/db_xref="taxon:4565"
/clone="wtls.pk005.p2"
/tissue_type="leaf"
/lab_host="DH10B"
/clone_1ib="wtls"
/note="Vector: pGEN-T Easy; Site 1: SmaI; R: band (susceptible) wheat leaves infected with Septoria tritici strain A, 24 hours after infection, subtracted w/ comparable uninfected leaves"

ORIGIN

Query Match 1.2%; Score 50.8; DB 14; Length 617;
Best Local Similarity 46.2%; Pred. No. 6.4;
Matches 165; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

3017 AGTAAGCAGAGCAGCAGACAACTTGAAGCCAAACAGTTGAACCGACTGCTAA 3076
161 AGAAGAA 220
3077 AACCAACAGAGTACCAAGTCCGCTCAAGAGAGAGAGAGAGAGAGAGAGAG 3136
221 AGAAGAA 280
3137 TACCTGCTGATCAAGCGCTGTTAAAGCATTAAGAGCAACAGCTGAAGTGC 3196
281 AGAAGAA 340
3197 TGAAGACAAAGAAAGTGAAGCAAAAGAAAGAAAGAAAGAAAGAAAGAAAG 3256
341 AGAAGAA 400
3257 TGATCCCTGCTGATTAAGAGCGCTGTTGCAATTGAAGCCGCACTTGAATGTC 3316
401 AGAAGAA 460
3317 CCCACAGCATCGGAAAGATCGTCTAGCTCAAGAGAGAGAGAGAGAGAGAGAG 3376
461 AGAA 520
3377 AAAAAA 3382
Db 521 AGAAGA 526

RESULT 13
B2231524/c 770 bp DNA linear GSS 12-OCT-2002
LOCUS
DEFINITION CH230-438K4.TV CHORI-230 Segment 2 Ratius norvegicus genomic clone
ACCESSION B2231524
VERSION B2231524.1 GI:23890065

KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 770)
AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K., Shwartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D., Riggs, P., de Jong, P. and Fraser, C.M.
TITLE Rat BAC End Sequences from Library CHORI-230 MboI segment
JOURNAL Unpublished (1999)
COMMENT Other GSSs: CH230-438K4.TJ
 Contact: Shaying Zhao
 Department of Bukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230 (http://www.chori.org/bacpac/rat230.htm). For BAC library availability, please contact Pieter de Jong (pdejong@tigr.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/or eting information.html). BAC end page: http://www.tigr.org/cdb/bac_ends/rat/bac_end_intro.html
 Plate: 438 row: K column: 4
 Seq primer: T7
 Class: BAC ends.

FEATURES
 Location/Qualifiers
 source 1..770
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SnHsd/MCw"
 /db_xref="taxon:10116"
 /clone="CH230-438K4"
 /sex="Female"
 /cell_type="Brain"
 /clone_1ib="CHORI-230 Segment 2"
 /note="Vector: pTARHAC1.3; Site 1: MboI; Site 2: MboI; CHORI-230 Rat (BN/SnHsd/MCw) BAC library produced by Pieter de Jong"

ORIGIN
 Query Match 1.2%; Score 50.2; DB 28; Length 770;
 Best Local Similarity 49.8%; Pred. No. 8.4; Indels 0; Gaps 0;
 Matches 127; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

2992 AAGCAGCAATTCACANTGATTAGTAAGACAGCAGCAGACGACATTAGAG 3051
 |||||
 712 AAG 653
 |||||
 3052 CCAACCAAGTTGACCGACTGCTTAAACCAACAGCGTACCGAAAGTCCGTTAAGA 3111
 |||||
 653 CAG 593
 |||||
 3112 GAGCAGCAGCAGCAGCGTTCTCTGATCCCTGCTGATCAAAAGCTGTTAAAGCATTA 3171
 |||||
 593 AAG 533
 |||||
 3172 AAGCCAAACAGCTGAATGACTCTGTAACACAAACAAAGTAAGGCAAAACAAACAA 3231
 |||||
 533 AAG 473
 |||||
 3232 TCCGCTCAAAAAGAG 3246
 |||||
 472 GAGAGAGAGAGAGAG 458
 |||||

RESULT 14
LOCUS CNS0177Y 1101 bp DNA linear GSS 26-Jul-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC

ACCESSION BACN37B08 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
VERSION AL108676
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyritroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1101)

REFERENCE 1
AUTHORS Direct Submission
TITLE Genoscope.
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage; BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.crs.fr
 - Web : www.genoscope.crs.fr)
 Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billand at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MEC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.

FEATURES
 Location/Qualifiers
 source 1..1101
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACN37B08"
 /clone_1ib="DrosBAC"
 /plasmid="pBelOBAC11"
 /note="end : T7"

ORIGIN
 Query Match 1.2%; Score 50; DB 29; Length 1101;
 Best Local Similarity 30.9%; Pred. No. 9;
 Matches 98; Conservative 83; Mismatches 132; Indels 4; Gaps 1;

777 AAGGGGAGACGGTTCGATGTTTATATATGCTGAAACAAACAAATGTTAAAT 836
 |||||
 750 AATGAGATTTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 809
 |||||
 837 AATGAGATTTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 896
 |||||
 810 AATGAGATTTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 869
 |||||
 897 CGCAATCTTATTTGATGAATTTTCAAGAGATTACATCACTTACACCCGA 956
 |||||
 870 AATGAGATTTTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 929
 |||||
 957 GCTGTAATGAG 1012
 |||||
 930 AATGAGATTTTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 989
 |||||
 1013 TCAGATATGAGAGATTAACATCAAGATTAATTAATTAATTAATTAATTAATTA 1072
 |||||
 990 AATGAGATTTTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1049
 |||||
 1073 GAAAGAGAGAGATTAAG 1089
 |||||
 1050 AATGAGATTTTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1066
 |||||

RESULT 15
LOCUS BK376338/c 1111 bp mRNA linear EST 08-MAY-2003
DEFINITION BK376338 Homo sapiens NEUROBLASTOMA COR 25-NORMALIZED Homo sapiens
ACCESSION BK376338
VERSION BK376338.1 GI:30446522
KEYWORDS EST.
SOURCE Homo sapiens (human)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2004, 22:15:39 ; Search time 1521 Seconds
(without alignments)
12063.095 Million cell updates/sec

Title: US-10-645-655-1

Perfect score: 4319
Sequence: 1 TCAGATGCTTAACTAAGT.....ACCTTTTATTCAGATTAAT 4319

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*
1: geneseqn1980s.*
2: geneseqn1980s.*
3: geneseqn2000s.*
4: geneseqn2000s.*
5: geneseqn2000s.*
6: geneseqn2000s.*
7: geneseqn2003as.*
8: geneseqn2003bs.*
9: geneseqn2003cs.*
10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4319	100.0	4319	2	AAT17215
2	4319	100.0	4319	8	ADA09342
3	3180.8	73.6	4305	8	ADA09349
4	3018.8	69.9	110000	2	AAT17203_02
5	2990.4	69.2	5245	8	ADA09353
6	2970.2	64.1	4822	8	ADA09355
7	2709.6	62.7	4605	8	ADA09351
8	1353.8	31.3	4350	7	ACA11657
9	1348	31.2	4350	2	AAZ12252
10	1297.8	30.0	4407	2	AAZ12253
11	1294.8	30.0	4404	7	ABZ28249
12	1294.8	30.0	4404	7	ACA0973
13	1252.6	29.0	4374	3	AAZ38918
14	1249.4	28.9	4374	2	AAZ12251
15	1249.4	28.9	4374	2	AAZ1302
16	1249.4	28.9	4374	3	AAZ1587
17	1249.4	28.9	4374	3	AAZ1587
18	1249.4	28.9	4374	3	AAZ1587
19	1249.4	28.9	4374	3	AAZ1587
20	1249.4	28.9	4374	3	AAZ1587
21	1234.8	28.6	4365	3	AAZ38917
22	969	22.4	969	2	AAZ88142
23	415.2	9.6	1312	3	AAZ81400

24	414	9.6	2991	2	AAZ12250	AAZ12250 Neisseria
25	202	4.7	5085	7	ACA34323	ACA34323 Prokaryot
26	202	4.7	110000	2	AAT17203_10	Continuation (11) o
27	189.2	4.4	5091	2	AAZ06164	AAZ06164 The iga g
28	174.6	4.0	30078	3	AAZ81520	AAZ81520 N. mening
29	174.6	4.0	110000	3	AAZ81490_07	Continuation (8) of
30	174.6	4.0	34980	3	AAZ1608	AAZ1608 Neisseria
31	168.4	3.9	891	3	AAZ1401	AAZ1401 N. mening
32	165	3.8	781	3	AAZ1402	AAZ1402 N. mening
33	164.4	3.8	5322	7	ACA41763	ACA41763 Prokaryot
34	155.2	3.6	4695	7	ABZ39057	ABZ39057 N. gonorr
35	155.2	3.6	4779	7	ACA41208	ACA41208 Prokaryot
36	135.2	3.1	4296	3	AAZ54326	AAZ54326 Neisseria
37	135.2	3.1	16526	3	AAZ1472	AAZ1472 N. mening
38	135.2	3.1	110000	3	AAZ1489_6	Continuation (7) of
39	135.2	3.1	172325	3	AAZ1613	AAZ1613 Neisseria
40	135.2	3.1	34980	3	AAZ1612	AAZ1612 Neisseria
41	126.8	2.9	4113	4	ABA89172	ABA89172 Escherich
42	126.8	2.9	48012	6	ABA79084	ABA79084 E. coli C
43	126.8	2.9	48254	4	ABA89141	ABA89141 Escherich
44	126.8	2.9	48345	4	ABA89142	ABA89142 Escherich
45	100.6	2.3	4128	4	ABA89994	ABA89994 Escherich

ALIGNMENTS

RESULT 1
AAT17215 standard; DNA, 4319 BP.
AC AAT17215;
DT 01-JUN-1996 (first entry)

Adhesion and penetration protein gene.

Adhesion and penetration protein; hap gene; promoter; terminator;
inverted repeat; stem-loop; protease; outer membrane protein;
beta-domain; secretion; recombinant vaccine; monoclonal antibody;
diagnostic; immunoassay; ds.

Haemophilus influenzae.

Key Location/Qualifiers

FT -35_signal

FT -10_signal

FT CDS

FT sig_peptide

FT mat_peptide

FT cat_peptide

FT mat_peptide

FT mat_peptide

FT terminator

FT repeat_region

FT stem_loop

FT MO9605856-A1.

PD	29-FEB-1996.
XX	
PF	16-AUG-1995; 95MO-US010661.
XX	
PR	25-AUG-1994; 94US-00296791.
XX	
PA	{UNIV } UNIV WASHINGTON.
PA	{STRD } UNIV LEIAND STANFORD JUNIOR.
XX	
PI	St Geme JW, Falkow S;
DR	
DR	WPI; 1996-151147/15.
XX	P-PSDB; AAR92768.
XX	
PT	Hemophilus adhesion and penetration protein and corresponding DNA - used
XX	to produce vaccines against H. influenzae infection.
XX	
S3	Claim 5; Fig 6; 105pp; English.

CC The sequence encodes a Haemophilus influenzae adhesion and penetration
CC protein. The sequence (hap gene) includes putative -10 and -30 sequences
CC and a putative rho-independent terminator 3' to the hap stop codon. The
CC terminator contains interrupted inverted repeats, with the potential for
CC forming a hairpin structure containing a loop of 3 bases and a stem of 8
CC bases, followed by a stretch rich in T residues. The gene product is
CC first synthesised as a preprotein, which is transported to the periplasm,
CC followed by insertion of the C-terminal beta-domain into the outer
CC membrane, possibly forming a pore, and export of the N-terminal fragment
CC through the outer membrane, followed by autoproteolytic cleavage and
CC secretion of the mature protease, leaving an outer membrane protein
CC fragment. The gene may be inserted in a vector and expressed in
CC recombinant host cells, for use as a recombinant vaccine. The gene
CC product may also be used in diagnostic monoclonal antibody production
XX
SQ Sequence 4319 BP; 1497 A; 776 C; 891 G; 1155 T; 0 U; 0 Other;

Query Match	100.0%;	Score 4319;	DB 2;	Length 4319;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 4319; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	1	CGAATAGCTTAACTAGTAATTTTAAATGAAAAATTAATCTATTAATTAATTAACATTA	60
Dp	1	TCATATGCTTTAACTAGTAATTTTAAATGAAAAATTAATCTATTAATTAATTAACATTA	60
Qy	61	TGAAAAAACTGTATTCGCTTAAATTTTAAACCGCTGCATTTCACTTAAGGATAGTAT	120
Dp	61	TGAAAAAACTGTATTCGCTTAAATTTTAAACCGCTGCATTTCACTTAAGGATAGTAT	120
Qy	121	CGCAAGCGGCGCTGTCACACTTAATTTTGGATAGTATCCAAATATTAACGATTTTG	180
Dp	121	CGCAAGCGGCGCTGTCACACTTAATTTTGGATAGTATCCAAATATTAACGATTTTG	180
Qy	181	CCGAGATAAAGGGAAGTTCAAGTTGGGGCTCCAAATATTAAGTTTATTAACAACAG	240
Dp	181	CCGAGATAAAGGGAAGTTCAAGTTGGGGCTCCAAATATTAAGTTTATTAACAACAG	240
Qy	241	GGCAATAGTTGGCACTCAATATGACAAAAGCCCGCATGTATTTTCTGTAGTGTAC	300
Dp	241	GGCAATAGTTGGCACTCAATATGACAAAAGCCCGCATGTATTTTCTGTAGTGTAC	300
Qy	301	GTAACGGCGTGCACGCTTGTTGAAATTCATATATTTGAGCGCTGCACATTAACGAG	360
Dp	301	GTAACGGCGTGCACGCTTGTTGAAATTCATATATTTGAGCGCTGCACATTAACGAG	360
Qy	361	GATATACAGTGTGATTTTGGTGTGACAGGGAAAACAACCCGATCAACATCGTTTACTT	420
Dp	361	GATATACAGTGTGATTTTGGTGTGACAGGGAAAACAACCCGATCAACATCGTTTACTT	420
Qy	421	ATAAGATTTGAAACGAAATAACTACAAAAAGATATTTATCATCTTATGAGACGATT	480
Dp	421	ATAAGATTTGAAACGAAATAACTACAAAAAGATATTTATCATCTTATGAGACGATT	480
Qy	481	ACGATATTCACGATTAATTAATTTGCTACAGAGCGCTCAATGATATAGCTTGA	540

Db 481 ACCGATATCCAGCATTTACATTAATTCGTTACAGAGCGGCTCCATTGATATGACTTCCA 540

Qy 541 AATGATATGCGAGTACTTATTCAGATAGAA CAATAATCCAGAA CGTGTCCGATCGCT 600

Db 541 AATATGATATGCGAGTACTTATTCAGATAGAA CAATAATCCAGAA CGTGTCCGATCGCT 600

Qy 601 CTGGAGCGCGAGTTTGGCGAATATGATCTAAACCAAAAGCGACAGCTTCCCGGTGCATATC 660

Db 601 CTGGAGCGCGAGTTTGGCGAATATGATCTAAACCAAAAGCGACAGCTTCCCGGTGCATATC 660

Qy 661 ATTATCTGACAGCTGCGCAATACACCAATGAGGTGAGACAGTATAGCATATTCGATAT 720

Db 661 ATTATCTGACAGCTGCGCAATACACCAATGAGGTGAGACAGTATAGCATATTCGATAT 720

Qy 721 TGGAGGCGCATGTTGGTAAAGCGGAGATATGGTCCATTACCGATTGCAAGCTCAAAAG 780

Db 721 TGGAGGCGCATGTTGGTAAAGCGGAGATATGGTCCATTACCGATTGCAAGCTCAAAAG 780

Qy 781 GGGACAGGTGCTCCGATGTTTATTAATGATCTGCAAAAA CAATAATGGTATATATATG 840

Db 781 GGGACAGGTGCTCCGATGTTTATTAATGATCTGCAAAAA CAATAATGGTATATATATG 840

Qy 841 GGAATATTAAGGCAAGCGCAACCTTTGAAGCGAAAGAAATGGGTTCAAATGGTTCGCA 900

Db 841 GGAATATTAAGGCAAGCGCAACCTTTGAAGCGAAAGAAATGGGTTCAAATGGTTCGCA 900

Qy 901 AATCTTATTTTGAATATTTTGCAGAGATTTTACATATACATCCCTTAACCCGCAAGCG 960

Db 901 AATCTTATTTTGAATATTTTGCAGAGATTTTACATATACATCCCTTAACCCGCAAGCG 960

Qy 961 GTATATGAGTATACCAATATATGTAATATGATATGGTCAAGGGTCTATATACCAAAAT 1020

Db 961 GTATATGAGTATACCAATATATGTAATATGATATGGTCAAGGGTCTATATACCAAAAT 1020

Qy 1021 CAGGAATACCAATATATATATATATGCTTACAGCAATATGATATTCCTTTGAAGAGA 1080

Db 1021 CAGGAATACCAATATATATATATATGCTTACAGCAATATGATATTCCTTTGAAGAGA 1080

Qy 1081 AGGATAAAGTTCATATCTAGATATACAGCACTATATATTTATCTCCACGTTTAAACA 1140

Db 1081 AGGATAAAGTTCATATCTAGATATACAGCACTATATATTTATCTCCACGTTTAAACA 1140

Qy 1141 ATGAGAAACGCTATATTTTATGATCAAAAACAGGATCATATCTTCGCACTTGACA 1200

Db 1141 ATGAGAAACGCTATATTTTATGATCAAAAACAGGATCATATCTTCGCACTTGACA 1200

Qy 1201 TTTAAACAAAGGGAGGTGGTCTTTATTTTGAAGGTATTTTACGATATCTCCAAATCTGA 1260

Db 1201 TTTAAACAAAGGGAGGTGGTCTTTATTTTGAAGGTATTTTACGATATCTCCAAATCTGA 1260

Qy 1261 ACCAAACTTGGCAAGGAGCTGAGCATACATGTAAAGTAAATAGCACCGTATCTTGGAAAG 1320

Db 1261 ACCAAACTTGGCAAGGAGCTGAGCATACATGTAAAGTAAATAGCACCGTATCTTGGAAAG 1320

Qy 1321 TTAATGGGTGGAGCACTGATGACCTTTCAAAATGGGTAAAGAGCAATGCAAGTTCAAG 1380

Db 1321 TTAATGGGTGGAGCACTGATGACCTTTCAAAATGGGTAAAGAGCAATGCAAGTTCAAG 1380

Qy 1381 CCAAAAGGGCAAAATAAAGGTTTCGATACGAGCTAGCGATGTAAAGTATTTTGGAGCAGC 1440

Db 1381 CCAAAAGGGCAAAATAAAGGTTTCGATACGAGCTAGCGATGTAAAGTATTTTGGAGCAGC 1440

Qy 1441 AAGCAGACGATCAAGGCACAAACAGCTTTTATGTAATATGGCTTATGAGGCAAG 1500

Db 1441 AAGCAGACGATCAAGGCACAAACAGCTTTTATGTAATATGGCTTATGAGGCAAG 1500

Qy 1501 GGAAGTCTCAATTAACGATGTAAACATTTGATATACGATACGATATTTTATTCGGCTTC 1560

Db 1501 GGAAGTCTCAATTAACGATGTAAACATTTGATATACGATACGATATTTTATTCGGCTTC 1560

Qy 1561 GTGGTGGTGGCTTAAGTCTTAAGGGGCAATTCATTAACCTTTAAACGATCCAAATACGG 1620

Db 1561 GTGGTGGTGGCTTAAGTCTTAAGGGGCAATTCATTAACCTTTAAACGATCCAAATACGG 1620


```

QY 3781 AAGAACAAAGCCGAAATTCATCGAAAAGCATTAATTAATGAGTATGCAAGTATC 3840
DB 3781 AAGAACAAAGCCGAAATTCATCGAAAAGCATTAATTAATGAGTATGCAAGTATC 3840
QY 3841 AGTCCGTTTAAAGGCAATTTGGCAGTATGAGTATGAGTATGAGTATGAGTATG 3900
DB 3841 AGTCCGTTTAAAGGCAATTTGGCAGTATGAGTATGAGTATGAGTATGAGTATG 3900
QY 3901 AACGTGAATTAATTCATCTAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3960
DB 3901 AACGTGAATTAATTCATCTAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3960
QY 3961 GGTATATATGTCGTCATTCAGTATGATTAATTAATTAATTAATTAATTAATTAAT 4020
DB 3961 GGTATATATGTCGTCATTCAGTATGATTAATTAATTAATTAATTAATTAATTAAT 4020
QY 4021 AGCCTTATTTCTGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 4080
DB 4021 AGCCTTATTTCTGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 4080
QY 4081 ATCTACGAGTGTGCAACACCATTTGACATTTGACATTTGACATTTGACATTTGAC 4140
DB 4081 ATCTACGAGTGTGCAACACCATTTGACATTTGACATTTGACATTTGACATTTGAC 4140
QY 4141 CAGAAATTTTCACTTCCAAATTCCTTTTATCTCAAAATCTCAAAATCTCAAAATCT 4200
DB 4141 CAGAAATTTTCACTTCCAAATTCCTTTTATCTCAAAATCTCAAAATCTCAAAATCT 4200
QY 4201 GCAGAACGCAAAATGAGGCGTGAATTTGGGCTATGTTGTTAAATTAATTAATTA 4260
DB 4201 GCAGAACGCAAAATGAGGCGTGAATTTGGGCTATGTTGTTAAATTAATTAATTA 4260
QY 4261 ATCGTTATTTGATTAACAAAGTGGTCAATTCAGATTCAGATTCAGATTCAGATTC 4319
DB 4261 ATCGTTATTTGATTAACAAAGTGGTCAATTCAGATTCAGATTCAGATTCAGATTC 4319

```

RESULT 2

AD09342
ID AD09342 standard; DNA; 4319 BP.

AC AD09342;

DT 06-NOV-2003 (first entry)

DE Haemophilus adherence and penetration protein (HAP) gene.

XX ds: Haemophilus adherence and penetration protein; HAP; gene;

KX antibacterial; immunostimulant; vaccine; Haemophilus influenza infection;

KX influenza.

OS Haemophilus influenzae.

XX US2003073166-A1.

XX 17-APR-2003.

XX 22-FEB-2002; 2002US-00080505.

XX 25-AUG-1994; 94US-00296791.

XX 20-APR-2001; 2001US-00839996.

XX (GENE/) GENE J W S.

XX Gene JWS;

XX WPI: 2003-567308/53.

XX P-PSDB; AD09343.

XX New recombinant Haemophilus adherence and penetration protein, useful as a

XX vaccine for prophylactic or therapeutic treatment of a Haemophilus

XX influenzae infections, in generating antibodies for passive immunization.

PS Example 1; Fig 6; 149pp; English.

XX The invention relates to a new recombinant Haemophilus adherence and
CC penetration protein (HAP) appearing as AD09348, AD09350, AD09351, and
CC AD09354 encoded by a nucleic acid appearing as AD09349, AD09351, and
CC AD09353, AD09355 and AD09357. Also included are a recombinant nucleic
CC acid that will hybridize under stringency conditions to the nucleic acid
CC detailed above, an expression vector comprising transcriptional and
CC translation regulatory nucleic acid operably linked to the HAP nucleic
CC acid, producing HAP by culturing a host cell transformed with the
CC expression vector; a composition comprising a pharmaceutical carrier and
CC HAP, inducing an immune response in a patient by administering the
CC composition, a composition comprising an antibody capable of binding to
CC HAP (or to a peptide selected from 35 peptide sequences appearing as
CC AD09359-AD09393), and an antibody capable of binding to the peptides.
CC The Haemophilus adherence and penetration (HAP) protein is useful as a
CC vaccine for prophylactic or therapeutic treatment of an H. influenzae
CC infection (e.g. influenza), in generating an immune response in a
CC patient, and in diagnosing the presence of H. influenzae infection in a
CC sample or patient, in passive immunisation treatments, and to purify or
CC separate HAP proteins or the H. influenzae organism from a sample. The
CC present sequence is the H. influenzae HAP gene.

SQ Sequence 4319 BP; 1497 A; 776 C; 891 G; 1155 T; 0 U; 0 Other;

Query Match 100.0%; Score 4319; DB 8; Length 4319;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TCAATAGTGTGTTACATGATTTTAAATGCAAAATTAATTAATTAATTAATTAATTA 60
DB 1 TCAATAGTGTGTTACATGATTTTAAATGCAAAATTAATTAATTAATTAATTAATTA 60
QY 61 TGAATAAAGCTGATTTGCTTAATTTTAAATGCAAAATTAATTAATTAATTAATTAAT 123
DB 61 TGAATAAAGCTGATTTGCTTAATTTTAAATGCAAAATTAATTAATTAATTAATTAAT 123
QY 121 CGAGACGTGGCTGTGACATTAATTTGGGATTTGATTAATTAATTAATTAATTAAT 180
DB 121 CGAGACGTGGCTGTGACATTAATTTGGGATTTGATTAATTAATTAATTAATTAAT 180
QY 181 CCGAGATTAAGGAGATTTGACATTTGGGCTCAAAATTAATTAATTAATTAATTAAT 240
DB 181 CCGAGATTAAGGAGATTTGACATTTGGGCTCAAAATTAATTAATTAATTAATTAAT 240
QY 241 GGCATTAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
DB 241 GGCATTAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
QY 301 GTAAAGCGGTGGCAGCTTGTGAAATTAATTAATTAATTAATTAATTAATTAAT 360
DB 301 GTAAAGCGGTGGCAGCTTGTGAAATTAATTAATTAATTAATTAATTAATTAAT 360
QY 361 GATATACAGTGTGATTTTGGTGGAGGGAACACCCGATCAATATGTTTACTT 420
DB 361 GATATACAGTGTGATTTTGGTGGAGGGAACACCCGATCAATATGTTTACTT 420
QY 421 ATAAAGTTGTAAAGCAAACTACTCAAAAAGATTAATTAATTAATTAATTAATTAAT 480
DB 421 ATAAAGTTGTAAAGCAAACTACTCAAAAAGATTAATTAATTAATTAATTAATTAAT 480
QY 481 ACCATTAATCAAGATTAATTAATTTGTTACAGAGCGGCTCAATATGATTCGA 540
DB 481 ACCATTAATCAAGATTAATTAATTTGTTACAGAGCGGCTCAATATGATTCGA 540
QY 541 ATATGAATGGGCACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
DB 541 ATATGAATGGGCACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
QY 601 CTGAGACGAGTTTGGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
DB 601 CTGAGACGAGTTTGGCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660

```

Oy	661	ATTATCTGACACTGCGCAATACACACATCAAGCGTGGACAGATTAATGATCTGATTT	720
Db	661	ATTATCTGACACTGCGCAATACACACATCAAGCGTGGACAGATTAATGATCTGATTT	720
Oy	721	TGGAGGCGCATCTTGTTAAAGCGGAGATATATGTCATTTACCGATTGCAAGCTCAAAAG	780
Db	721	TGGAGGCGCATCTTGTTAAAGCGGAGATATATGTCATTTACCGATTGCAAGCTCAAAAG	780
Oy	781	GGGACAGTGGTCTCCGATCTTTATATATGATCTCTGTAATAAAACAAAATGGTAAATATG	840
Db	781	GGGACAGTGGTCTCCGATCTTTATATATGATCTCTGTAATAAAACAAAATGGTAAATATG	840
Oy	841	GGAATATTACGGGAGAGCAACCCCTTTGAAAGCAAAAAGGGCTTCAATTTGGTTCGA	900
Db	841	GGAATATTACGGGAGAGCAACCCCTTTGAAAGCAAAAAGGGCTTCAATTTGGTTCGA	900
Oy	901	AATCTTAATTTTGAAGAAATTTTTCGAAAGATTTACATTCATCACTTTACACCGAGCTG	960
Db	901	AATCTTAATTTTGAAGAAATTTTTCGAAAGATTTACATTCATCACTTTACACCGAGCTG	960
Oy	961	GTATCGAGGTGACACAAATTAGTGAATATGATATATGTCAGGGGTCTAATPACTCAGAAAT	1020
Db	961	GTATCGAGGTGACACAAATTAGTGAATATGATATATGTCAGGGGTCTAATPACTCAGAAAT	1020
Oy	1021	CAGCAATACCATCGAATATTAATTAATTCGTTAGCAATATATGATTTACCTTTGAAAGAG	1080
Db	1021	CAGCAATACCATCGAATATTAATTAATTCGTTAGCAATATATGATTTACCTTTGAAAGAG	1080
Oy	1081	AGGATTAATCTTCATATACCTCGATATATGACGGACCTTAATTTATCTTCACAGTTTAAAC	1140
Db	1081	AGGATTAATCTTCATATACCTCGATATATGACGGACCTTAATTTATCTTCACAGTTTAAAC	1140
Oy	1141	ATGAGAGAAACGCTATATTTATATGATCAAAAACAGAGATCTTAATCTTCGATCTGACA	1200
Db	1141	ATGAGAGAAACGCTATATTTATATGATCAAAAACAGAGATCTTAATCTTCGATCTGACA	1200
Oy	1201	TTTACCAAGGGGCGGGGTCTTTATTTTGAAGGTAAATTTTACAGTATCTCCAAATCTTA	1260
Db	1201	TTTACCAAGGGGCGGGGTCTTTATTTTGAAGGTAAATTTTACAGTATCTCCAAATCTTA	1260
Oy	1261	ACCAAACTTGGCAGAGAGCTGCGCATATCATGTATGATTAAGATACCGTATCTGGAAAG	1320
Db	1261	ACCAAACTTGGCAGAGAGCTGCGCATATCATGTATGATTAAGATACCGTATCTGGAAAG	1320
Oy	1321	TAAATGCGCTGGAAATCATGATCGACTTCTMAAATTTGTATTAAGAAACATGCACTTCAG	1380
Db	1321	TAAATGCGCTGGAAATCATGATCGACTTCTMAAATTTGTATTAAGAAACATGCACTTCAG	1380
Oy	1381	CGAAAGGGGAAATTAAGATGTGATATCAGCGTGAAGCGATATGATCATTTTGGAGAGAG	1440
Db	1381	CGAAAGGGGAAATTAAGATGTGATATCAGCGTGAAGCGATATGATCATTTTGGAGAGAG	1440
Oy	1441	AGGCAAGCATCAGAGGCAACAAACAAGCTTTAGTGAATATGCGTTGTATGACGGCAGAG	1500
Db	1441	AGGCAAGCATCAGAGGCAACAAACAAGCTTTAGTGAATATGCGTTGTATGACGGCAGAG	1500
Oy	1501	GGAATCTTCATTAATAAGATATTAACAATTTGATACCGATTAATTTTATTTTCGGCTTTC	1560
Db	1501	GGAATCTTCATTAATAAGATATTAACAATTTGATACCGATTAATTTTATTTTCGGCTTTC	1560
Oy	1561	GTGATGGTGGCTTATGATCTTTAAGGGCACTTCACTTTAAACCTTTAAACGATATCCAAAATACG	1620
Db	1561	GTGATGGTGGCTTATGATCTTTAAGGGCACTTCACTTTAAACCTTTAAACGATATCCAAAATACG	1620
Oy	1621	ACGAGGGGCGAATGATTTGTAACCTAATATCACTCAAGCGGTAAATATGCACTTATCTG	1680
Db	1621	ACGAGGGGCGAATGATTTGTAACCTAATATCACTCAAGCGGTAAATATGCACTTATCTG	1680
Oy	1681	GGAAACGAAAGCAATGTTCTACCTTAAGGAAATTAATTTATTAATCACTGATATACGAAAG	1740
Db	1681	GGAAACGAAAGCAATGTTCTACCTTAAGGAAATTAATTTATTAATCACTGATATACGAAAG	1740

QY	1741	AAATTCCTCAAAACGGTGTGGTGGCAAAACGATTAATAATTAACAATGGCGCATTA	1800
Db	1741	AAATTCCTCAAAACGGTGTGGTGGCAAAACGATTAATAATTAACAATGGCGCATTA	1800
QY	1801	ACCTTATTATTAACCAACCAAGAGATGTACTTCTACTTTCAAGTGGTACAAAT	1860
Db	1801	ACCTTATTATTAACCAACCAAGAGATGTACTTCTACTTTCAAGTGGTACAAAT	1860
QY	1861	TAAAGGGCGATTTATCCCAACAAAGTAACTTTTTCAGCGGTAGACCGACACGGC	1920
Db	1861	TAAAGGGCGATTTATCCCAACAAAGTAACTTTTTCAGCGGTAGACCGACACGGC	1920
QY	1921	ACGCTACAAATCATTTAATAAGTTGGTGGCAAAATGSAAGGTATACCAAGGCGAA	1980
Db	1921	ACGCTACAAATCATTTAATAAGTTGGTGGCAAAATGSAAGGTATACCAAGGCGAA	1980
QY	1981	TTGTGTGGGATCAACGATTTGGATCAACCGTACATTTAAAGCTGAAACTTCCAAATTAAG	2040
Db	1981	TTGTGTGGGATCAACGATTTGGATCAACCGTACATTTAAAGCTGAAACTTCCAAATTAAG	2040
QY	2041	GCGGAAGTGCCTGTGTTCTGGCAATGTTCTTCAATGAGGAAATTTGCAACGTACGA	2100
Db	2041	GCGGAAGTGCCTGTGTTCTGGCAATGTTCTTCAATGAGGAAATTTGCAACGTACGA	2100
QY	2101	ATAATGCAAAATGGCAACATTTGTGTGTGGCAAAATCAACAAATACATTTGCAAGGTT	2160
Db	2101	ATAATGCAAAATGGCAACATTTGTGTGTGGCAAAATCAACAAATACATTTGCAAGGTT	2160
QY	2161	CAGATTGGACAGATTTAAGCATTTGTCAAAAGTGGATTTAACCGTACAAAGTTATTA	2220
Db	2161	CAGATTGGACAGATTTAAGCATTTGTCAAAAGTGGATTTAACCGTACAAAGTTATTA	2220
QY	2221	ATTCTATACCAAAACCAATCAATGAGCTCTATTATTAACGTATACGTAAACGAGCGA	2280
Db	2221	ATTCTATACCAAAACCAATCAATGAGCTCTATTATTAACGTATACGTAAACGAGCGA	2280
QY	2281	ATGTATTAAGTTTGGCAAACTTAAATGGCAATGTGACCTTTAACAATCACGCGCAATTA	2340
Db	2281	ATGTATTAAGTTTGGCAAACTTAAATGGCAATGTGACCTTTAACAATCACGCGCAATTA	2340
QY	2341	CATTAGCAACAATGCCCAACCTAAGGCATATTGCACTTCCGCAATTCACATGCGAA	2400
Db	2341	CATTAGCAACAATGCCCAACCTAAGGCATATTGCACTTCCGCAATTCACATGCGAA	2400
QY	2401	CGGTGATATATGCAAACTTGAACGTATATGTGCATTTAAGCATATCAGTCAATTTCTT	2460
Db	2401	CGGTGATATATGCAAACTTGAACGTATATGTGCATTTAAGCATATCAGTCAATTTCTT	2460
QY	2461	TAAAAAACAGCAATTTTGGCAACAAATTCAGAGGACAAAGGCAACAAGTACGTGTG	2520
Db	2461	TAAAAAACAGCAATTTTGGCAACAAATTCAGAGGACAAAGGCAACAAGTACGTGTG	2520
QY	2521	AAATTCGATGGAGCATGCTCAGCATCTACATCTGACGATTTTAAACGTAAATACAA	2580
Db	2521	AAATTCGATGGAGCATGCTCAGCATCTACATCTGACGATTTTAAACGTAAATACAA	2580
QY	2581	GTACGATCAACGTTAAATCAGCTTATTCAGCTAGCTCAACAAATACGCAAGTGCCTT	2640
Db	2581	GTACGATCAACGTTAAATCAGCTTATTCAGCTAGCTCAACAAATACGCAAGTGCCTT	2640
QY	2641	CATTAAACAGCAAAACAGGCAACATGCGACGAACATGTTCAACATTAAGCATTA	2700
Db	2641	CATTAAACAGCAAAACAGGCAACATGCGACGAACATGTTCAACATTAAGCATTA	2700
QY	2701	ATGGTAAATTAAGTGGGCAAGGCAATTCGAATTAATGATCTTTATTTGGCTATTA	2760
Db	2701	ATGGTAAATTAAGTGGGCAAGGCAATTCGAATTAATGATCTTTATTTGGCTATTA	2760
QY	2761	GGCATTAATTAATTAATTCGAATGACGCTGAGGCGATTAACATTAATCTGTGCGACA	2820
Db	2761	GGCATTAATTAATTAATTCGAATGACGCTGAGGCGATTAACATTAATCTGTGCGACA	2820
QY	2821	CAGGCAAAAGAACCCGAAACCTTGAAGCATTAACCTTGTGTAAAGCAAAAGATACAAAC	2880


```

Db      2821  CAGGCAAGACCCCAAAACCTTGACAAATTAATTTGTGAAACAAAGATTAATCAAC 2880
Qy      2881  CGTATCAGTAGTCAATTTACTTTGAAAAATGACACAGCTGATGAGGTGATTC 2940
Db      2881  CGTATCAGTAGTCAATTTACTTTGAAAAATGACAGGTGATGAGGTGATTC 2940
Qy      2941  GTTATTAATTAAGTAAGATGATGAGCAATTCGCTTGCATTAACCAATTAAGAGAG 3000
Db      2941  GTTATTAATTAAGTAAGATGATGAGCAATTCGCTTGCATTAACCAATTAAGAGAG 3000
Qy      3001  AATTGCAATATGATTTAGTAAGACAGACCAAGCAACCAATTAAGAGCAAG 3060
Db      3001  AATTGCAATATGATTTAGTAAGACAGACCAAGCAACCAATTAAGAGCAAG 3060
Qy      3061  TTGACCGATGCTTAAACAAACAGGTGACCAAAAGTGGCTGAATAAGACAGCA 3120
Db      3061  TTGACCGATGCTTAAACAAACAGGTGACCAAAAGTGGCTGAATAAGACAGCA 3120
Qy      3121  GAGCAGCGTTCCTGATACCTGCTGATCAAAAGCTGTTAAAGCATTAAGAGCAAC 3180
Db      3121  GAGCAGCGTTCCTGATACCTGCTGATCAAAAGCTGTTAAAGCATTAAGAGCAAC 3180
Qy      3181  AAGTGAATCTGACCTGTAACCAAAAGTAAGCAAAACCAAAAGTGGCTGCA 3240
Db      3181  AAGTGAATCTGACCTGTAACCAAAAGTAAGCAAAACCAAAAGTGGCTGCA 3240
Qy      3241  AAAAGCAGTGTTCCTGATACCTGCTGATCAAAAGCTGTTAAAGCAAGCCGAC 3300
Db      3241  AAAAGCAGTGTTCCTGATACCTGCTGATCAAAAGCTGTTAAAGCAAGCCGAC 3300
Qy      3301  TTGAGTTATTAATGATCCCAACAGCAATGCAAAAGTGGCTGATCAAGAGAGCG 3360
Db      3301  TTGAGTTATTAATGATCCCAACAGCAATGCAAAAGTGGCTGATCAAGAGAGCG 3360
Qy      3361  AAAAACAAGCAAAACCAAAAGTGGCTGATCAAGAGAGAGAGAGAGAGAGAGAG 3420
Db      3361  AAAAACAAGCAAAACCAAAAGTGGCTGATCAAGAGAGAGAGAGAGAGAGAGAG 3420
Qy      3421  TATCTGCAAGTAATGATGATGCTTCTGCTGATCAAGAGAGAGAGAGAGAGAG 3480
Db      3421  TATCTGCAAGTAATGATGATGCTTCTGCTGATCAAGAGAGAGAGAGAGAGAG 3480
Qy      3481  ATCAAGCAATTCGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540
Db      3481  ATCAAGCAATTCGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540
Qy      3541  ATGCTTCGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
Db      3541  ATGCTTCGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
Qy      3601  CTTAGCTATGAGCAATTTGGGCAATTTTTCGATAGCCGTTGAGTAATACCTTG 3660
Db      3601  CTTAGCTATGAGCAATTTGGGCAATTTTTCGATAGCCGTTGAGTAATACCTTG 3660
Qy      3661  ATGAACAGTTAAATTAATCAACGCAATTAATGATGATGATGATGATGATGAT 3720
Db      3661  ATGAACAGTTAAATTAATCAACGCAATTAATGATGATGATGATGATGATGAT 3720
Qy      3721  GGGGCGATTAAGATTTGCTTAAGCTGGAACGGGAATTAAGTGGCAATTAATGCTG 3780
Db      3721  GGGGCGATTAAGATTTGCTTAAGCTGGAACGGGAATTAAGTGGCAATTAATGCTG 3780
Qy      3781  AAGAACAAGCCGAAATTAATCAAGCAATTAATGATGATGATGATGATGATGAT 3840
Db      3781  AAGAACAAGCCGAAATTAATCAAGCAATTAATGATGATGATGATGATGATGAT 3840
Qy      3841  AGTTCGTTTGGGCAATTTGGGCAATTTGAGCTTATTTGAGATTAATGCTATTTAT 3900
Db      3841  AGTTCGTTTGGGCAATTTGGGCAATTTGAGCTTATTTGAGATTAATGCTATTTAT 3900
Qy      3901  AACGTGAATTAATCAATGATGAGAGAGTGAAGTGAAGAGCCGATGATTTATTC 3960

```

```

Db      3901  AACGTGAATTAATCAATGATGAGAGAGTGAAGTGAAGAGAGAGAGAGAGAGAG 3960
Qy      3961  GCATTAATGCTGATGATGATTAATCAATTAATGATGATGATGATGATGATGAT 4020
Db      3961  GCATTAATGCTGATGATGATTAATCAATTAATGATGATGATGATGATGATGAT 4020
Qy      4021  AGCTTATTTCTGCTCAATTAATGATGATGATGATGATGATGATGATGATGAT 4080
Db      4021  AGCTTATTTCTGCTCAATTAATGATGATGATGATGATGATGATGATGATGAT 4080
Qy      4081  ATCTGACGCTGCAACCAATTTGGGCAATTTGAGCTTATTTGAGATTAATGCT 4140
Db      4081  ATCTGACGCTGCAACCAATTTGGGCAATTTGAGCTTATTTGAGATTAATGCT 4140
Qy      4141  CGAAATTTTCAATTTCAATTTCCGCTTATTAATGATGATGATGATGATGATGAT 4200
Db      4141  CGAAATTTTCAATTTCAATTTCCGCTTATTAATGATGATGATGATGATGATGAT 4200
Qy      4201  GCAACAGCAAAATGAGGCTGATGATGATGATGATGATGATGATGATGATGAT 4260
Db      4201  GCAACAGCAAAATGAGGCTGATGATGATGATGATGATGATGATGATGATGAT 4260
Qy      4261  ATGCTTATTAATTAACAGAGTGGCTGATGATGATGATGATGATGATGATGAT 4319
Db      4261  ATGCTTATTAATTAACAGAGTGGCTGATGATGATGATGATGATGATGATGAT 4319

```

RESULT 3

ADA09349

ID ADA09349 standard; DNA; 4305 BP.

ADA09349;

06-NOV-2003 (first entry)

Haemophilus HAP protein gene from strain 11.

ds; Haemophilus adherence and penetration protein; HAP; gene;

antibacterial; immunostimulant; vaccine; Haemophilus influenza infection;

Haemophilus influenzae; strain 11.

US2003073166-A1.

17-APR-2003.

22-FEB-2002; 2002US-03080505.

25-AUG-1994; 94US-00296791.

20-APR-2001; 2001US-00839996.

(GENE/) GENE J W S.

Gene JWS;

MPI; 2003-567308/53.

P-PsDB; ADA09350.

New recombinant Haemophilus adhesion and penetration protein, useful as a

vaccine for prophylactic or therapeutic treatment of a Haemophilus

influenzae infections, in generating antibodies for passive immunization.

Claim 1; Fig 16; 149P; English.

The invention relates to a new recombinant Haemophilus adhesion and

penetration protein (HAP) appearing as ADA09348, ADA09350, ADA09352 and

ADA09354 encoded by a nucleic acid appearing as ADA09349, ADA09351,

ADA09353, ADA09355 and ADA09357. Also included are a recombinant nucleic

acid that will hybridize under stringency conditions to the nucleic acid

described above, an expression vector comprising transcriptional and

translation regulatory nucleic acid operably linked to the HAP nucleic

acid, producing HAP, by culturing a host cell transformed with the

CC expression vector, a composition comprising a pharmaceutical carrier and
 CC HAP, inducing an immune response in a patient by administering the
 CC composition, a composition comprising an antibody capable of binding to
 CC HAP (or to a peptide selected from 35 peptide sequences appearing as
 CC AD09359-AD09393), and an antibody capable of binding to the peptides.
 CC The Haemophilus adhesion and penetration (HAP) protein is useful as a
 CC vaccine for prophylactic or therapeutic treatment of an H. influenzae
 CC infection (e.g. influenza), in generating an immune response in a
 CC patient, and in generating antibodies to the HAP protein. The antibodies
 CC may be used to diagnose the presence of H. influenzae infection in a
 CC sample or patient, in passive immunisation treatments, and to purify or
 CC separate HAP proteins or the H. influenzae organism from a sample. The
 CC present sequence is a HAP gene from a strain of H. influenzae.

XX Sequence 4305 BP; 1490 A; 762 C; 902 G; 1150 T; 0 U; 1 Other;

Query Match 73.6%; Score 3180.8; DB 8; Length 4305;
 Best Local Similarity 85.7%; Pred. No. 0;
 Matches 3704; Conservative 0; Mismatches 463; Indels 156; Gaps 9;

QY 60 ATGAAAAAAGCTGATTTCTGCTTAATTTTACCGCTTGCAATTCATTAGGATGTA 119
 DB 1 ATGAAAAAAGCTGATTTCTGCTTAATTTTACCGCTTGCAATTCATTAGGATGTA 60
 QY 120 TCGAAGCGTGGCTGCTGACATTTATTTGGATTGATTCGAAATTTATCGATTTT 179
 DB 61 TCGAAGCGTGGCTGCTGACATTTATTTGGATTGATTCGAAATTTATCGATTTT 120
 QY 180 GCGGAGATTAAGAGGAGTTCACAGTGGGCGCTCAAAATATTAAGTTTATACAAACA 239
 DB 121 GCGGAGATTAAGAGGAGTTCACAGTGGGCGCTCAAAATATTAAGTTTATACAAACA 180
 QY 240 GGGCAATTAAGTGGCAATCATGACAAAGCCCGATGATTTTCTGATGTC 299
 DB 181 GGGCAATTAAGTGGCAATCATGACAAAGCCCGATGATTTTCTGATGTC 240
 QY 300 CGTAAAGCGCTGGCAAGCTTGGTGAATCAATATTTGAGCGTGGCAATACGTA 359
 DB 241 AGAATGAGAGTTCCTGCTTGTAGTGGCAATGATTTGAGTGGCAATATGTA 300
 QY 360 GGAATATCAGATTTGATTTGTCGAGGAGAAACAACCCGATTAACATGCTTTACT 419
 DB 301 GGGTATACAGATTTGATTTGTCGAGGAGAAACAACCCGATTAACATGCTTTACT 360
 QY 420 TATAGATTTGTAAGAAATGACAAATCAAAAAAGATTAATTAACCTTATGAGAGCAT 479
 DB 361 TATAGATTTGTAAGAAATGACAAATCAAAAAAGATTAATTAACCTTATGAGAGCAT 420
 QY 480 TACCAATATCAGATTTGATTTGTCGAGGAGAAACAACCCGATTAACATGCTTTACT 539
 DB 421 TACCAATATCAGATTTGATTTGTCGAGGAGAAACAACCCGATTAACATGCTTTACT 480
 QY 540 AATATGATTTGAGAGTTCATTTGTCGAGGAGAAACAACCCGATTAACATGCTTTACT 599
 DB 481 AATATGATTTGAGAGTTCATTTGTCGAGGAGAAACAACCCGATTAACATGCTTTACT 540
 QY 600 TCTGAGAGGATTTGTCGAGGAGAAATGATCAAGCAAGGCGACCAAGTTCCCGTGCATAT 659
 DB 541 TCTGAGAGGATTTGTCGAGGAGAAATGATCAAGCAAGGCGACCAAGTTCTTATCTGAGAGAT 600
 QY 660 CATTAATGACAGCTGGCAATACACATATGAGCGTGGAGAGGATATGATTTGAT 719
 DB 601 AATTAATTAAGAGTGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
 QY 720 TTGGAGAGGATTTGTCGAGGAGAAATGATCAAGCAAGGCGACCAAGTTCCCGTGCATAT 779
 DB 661 GTGAGAGGATTTGTCGAGGAGAAATGATCAAGCAAGGCGACCAAGTTCTTATCTGAGAGAT 720
 QY 780 GGGGACAGTGGTCTCCGATTTATTTATGATGCTGAAGAAACAAATATGTTATAT 839
 DB 721 GGGGACAGTGGTCTCCGATTTATTTATGATGCTGAAGAAACAAATATGTTATAT 780
 QY 840 GGGATTTATGAGGAGGACACCTTTTGAAGGCAAGAAATGGGTTTCAATTTGCTGC 899

DB 781 GGGATTTGCGGGGGGCAACCTTTACAGGCAATTCGATGAGTTCATTAAGCCCT 840
 QY 900 AATCTTAATTTGATGAAATTTTCCAAAGAGATTTGATATCATCTTACACCGAGCT 959
 DB 841 AATCTTAATTTGATGAAATTTTCCAAAGAGATTTGATATCATCTTACACCGAGCT 900
 QY 960 GGTATGAGTGTACCAATATGAGTGAATGATTAAGTCAAGGCTTATCACTAGAAA 1019
 DB 901 GGTATGAGTGTATACATTTCTAAAGTGAATGATTAAGTCAAGGCTTATCACTAGAAA 960
 QY 1020 TCGAATTCATCAGAAATTAATTAATTAAGTGAATGATTAAGTCAAGGCTTATCACTAGAAA 1079
 DB 961 ATTGGAAGAAC-----TCGTGAATTAACCTTTACGAAACAAATTAATAAATGAAGAT 1014
 QY 1080 AAGGATTAAGTCAATATCTGATATTAAGAGAGCACTTAATTTATCTCCAGCTTAAC 1139
 DB 1015 AAGGATTAAGTCAATATCTGATATTAAGAGAGCACTTAATTTATCTCCAGCTTAAC 1074
 QY 1140 AATGAGAAAGCGCTATTTATGATCAAGAAACAGAGATCATTAATCTGATGAC 1199
 DB 1075 AATGAGAAAGATTTATTTGAGATGAAGATTAAGATCATTAATCTGATGAC 1134
 QY 1200 ATTAACAAAGGCGGCTGCTTATTTGAGGCTTAATTTACGATCTCAATTTCT 1259
 DB 1135 ATGATCAAGGCTGAGCGCTTGTATTTGAGGAGGATTTATGTTCCCTACCAAA 1134
 QY 1260 AACCAACCTTGGCAAGAGCTGCAATCATGATTAAGGAAATTAAGAGCTTACTGAAA 1319
 DB 1195 AATGAAGGATTAAGGAGCTGCAATCATGATTAAGGAAATTAAGAGCTTACTGAAA 1254
 QY 1320 GTAAATGAGGAGCAATCATGATCACTTTCTAAATTTGTAAGGAACTTGCCTCA 1379
 DB 1255 GTAAATGAGGAGCAATCATGATCACTTTCTAAATTTGTAAGGAACTTGCCTCA 1314
 QY 1380 GCCAAAGGGAATTAAGGCTGATCAAGCTGAGGAGTGAAGTCAATTTGAGAGAG 1439
 DB 1315 GCCAAAGGGAATTAAGGCTGATCAAGCTGAGGAGTGAAGTCAATTTGAGAGAG 1374
 QY 1440 CAGGAGAGATCAAGGCAACAAACAGCTTTAGTGAATTTGGCTGCTGAGCGAG 1499
 DB 1375 CAGGAGAGATCAAGGCAACAAACAGCTTTAGTGAATTTGGCTGCTGAGCGAG 1434
 QY 1500 GGGACTGTTCAATTAAGATGATTAACAAATTTGATACGATTAATTTTATTTGCTT 1559
 DB 1435 GGGACTGTTCAATTAAGATGATTAACAAATTTGATACGATTAATTTTATTTGCTT 1494
 QY 1560 CGTGGTGGTGGTGAATCTTAAGGAGATTTCAATTAATTTTAAAGTCAATTAAG 1619
 DB 1495 CGTGGTGGTGGTGAATCTTAAGGAGATTTCAATTAATTTTAAAGTCAATTAAG 1554
 QY 1620 GAGGAGGAGGAGATTTGAGCAATTAACATTAACCTTAAGCGCTATGCTATTAAT 1679
 DB 1555 GAGGAGGAGGAGATTTGAGCAATTAACATTAACCTTAAGCGCTATGCTATTAAT 1614
 QY 1680 GGGAGCAAGCAT--TGTTCACTTAATGAAATTAATTAATTAATTAATTAATTAAT 1736
 DB 1615 GGGAGCAAGCATTAATGCTCACTTAATTAATTAATTAATTAATTAATTAATTAAT 1674
 QY 1737 AAGGAATTTGCCCAAGGCTTGGTGGGAGAAACAGATTAATTAATTAATTAATTAAT 1796
 DB 1675 AAGGAATTTGCCCAAGGCTTGGTGGGAGAAACAGATTAATTAATTAATTAATTAAT 1734
 QY 1797 TTAACCTTATTTAAACCAACCAAGATGATTAATTAATTAATTAATTAATTAATTAAT 1856
 DB 1735 TTAACCTTATTTAAACCAACCAAGATGATTAATTAATTAATTAATTAATTAATTAAT 1794
 QY 1857 AATTAAGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1916
 DB 1795 AATTAAGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1854
 QY 1917 CCGGAGGCTTACATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1976

Db 1855 CCCCACCCCTACATCATTTAGACAAAGTGGTGAAGATGAGAGGTATCCCAAGAC 1914
 Qy 1977 GAATGTGTGGATACGATTTGATGACCCGATCACTTAAAGTGAAGAACTCCAAAT 2036
 Db 1915 GAAATGTGTGGATACGATTTGATGACCCGATCACTTAAAGTGAAGAACTCCAAAT 1974
 Qy 2037 AAAAGCGAAGTGGGTGGTTCCTGGCATGTTCCTCAATTGAGGAAATGGAAGTC 2096
 Db 1975 AAAAGCGAAGTGGGTGGTTCCTGGCATGTTCCTCAATTGAGGAAATGGAAGTC 2034
 Qy 2097 AGCAATATGCAAAATGCAATTTGGTGTGGCAATTCACAAATTCATTTGACG 2156
 Db 2035 AGCAATATGCAAAATGCAATTTGGTGTGGCAATTCACAAATTCATTTGACG 2094
 Qy 2157 GGTTCAGATTGACAGATTAAGCATGTCGCAAAAGTGAATTAAACGATCAAAAGT 2216
 Db 2095 GGTTCAGATTGACAGATTAAGCATGTCGCAAAAGTGAATTAAACGATCAAAAGT 2154
 Qy 2217 ATTAATTCATACCAAAACCAAAATGATGCTCATTAATTAATGATGATGACAG 2276
 Db 2155 ATTAATTCATACCAAAACCAAAATGATGCTCATTAATTAATGATGATGACAG 2214
 Qy 2277 GCGAATCTTAAGTTCGCAAACTTAATGCAATGTCATTTAAACAAATGACGCA 2336
 Db 2215 GCGAATCTTAAGTTCGCAAACTTAATGCAATGTCATTTAAATATGATGCA 2274
 Qy 2337 TTACATTAGCAATGATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2396
 Db 2275 TTACATTAGCAATGATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2234
 Qy 2397 ----- 2396
 Db 2335 GCAAGGTAACTATGATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2394
 Qy 2397 -----GCA 2399
 Db 2395 ACATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2454
 Qy 2400 ACATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2459
 Db 2455 ACATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2414
 Qy 2460 TTAATAAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2519
 Db 2515 TTAATAAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2574
 Qy 2520 GAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2579
 Db 2575 GAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2634
 Qy 2580 AGTACGATCAGTAAATTCAGTTTACGTTAGCTCAAACTAATGCAATGCAATGCAAT 2636
 Db 2635 AGTACGATCAGTAAATTCAGTTTACGTTAGCTCAAACTAATGCAATGCAATGCAAT 2634
 Qy 2637 GGTTCATTAAGACGAAACCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2696
 Db 2695 GGTTCATTAAGACGAAACCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2754
 Qy 2697 GGTTCATTAAGACGAAACCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2756
 Db 2755 GGTTCATTAAGACGAAACCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2814
 Qy 2757 AAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2816
 Db 2815 AAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2874
 Qy 2817 AAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2876
 Db 2875 AAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2934
 Qy 2877 CAAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2936
 Db 2935 CAAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2994

Qy 2937 TTACGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2996
 Db 2995 TTACGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3054
 Qy 2997 CAGAAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3056
 Db 3055 CAGAAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3114
 Qy 3057 CAGAAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3116
 Db 3115 CAGAAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3168
 Qy 3117 GCGAGAGAGGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3176
 Db 3169 ---AGAGGAGGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3225
 Qy 3177 AAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3236
 Db 3226 AAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3285
 Qy 3237 TCAAAAAGG-----AGTTCCTGATGATGATGATGATGATGATGATGATGATGAT 3284
 Db 3286 TCAAAAAGG-----AGTTCCTGATGATGATGATGATGATGATGATGATGATGAT 3345
 Qy 3285 GCAATTAAGAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3344
 Db 3346 GCAATTAAGAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3405
 Qy 3345 GCTCAA---GAAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3401
 Db 3406 GCTCAA---GAAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3465
 Qy 3402 AAATGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3461
 Db 3466 AAATGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3525
 Qy 3462 TTACGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3521
 Db 3526 TTACGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3585
 Qy 3522 AAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3581
 Db 3586 AAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3642
 Qy 3582 CAAATTAAGAGGAG 3641
 Db 3643 CAAATTAAGAGGAG 3702
 Qy 3642 GGTTCATTAAGACGAAACCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3701
 Db 3703 GGTTCATTAAGACGAAACCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3762
 Qy 3702 GGTTCATTAAGACGAAACCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3761
 Db 3763 GGTTCATTAAGACGAAACCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3822
 Qy 3762 AGTACGATCAGTAAATTCAGTTTACGTTAGCTCAAACTAATGCAATGCAATGCAAT 3821
 Db 3823 AGTACGATCAGTAAATTCAGTTTACGTTAGCTCAAACTAATGCAATGCAATGCAAT 3882
 Qy 3822 GGTTCATTAAGACGAAACCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3881
 Db 3883 GGTTCATTAAGACGAAACCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3942
 Qy 3882 GGTTCATTAAGACGAAACCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3941
 Db 3943 GGTTCATTAAGACGAAACCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 4002
 Qy 3942 CTAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4001
 Db 4003 CTAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4062

QY 1392 AATAAGGTTGATCAGCGTACGGATGTAAGTCAATTTTGAGCGACGAGACGAT 1451
 DB 77147 AATTAGGAAGTTAAGTGGCGGTACGGCAAGTCACTTACATCAACAAGACATGAA 77088
 QY 1452 CAAGCGACAAACAAACCTTTAGTGAATTTGGCTGTGTTACGGCAGAGGACCTTCA 1511
 DB 77087 GCGGGTCAAAAACAACTTTCAAGAAAGTTGGATGTGAACGGTTCAGACAACTTTCA 77028
 QY 1512 TTTAAACGATTAACAAATTTGATACCGATTAATTTTGGCTTGGTGGTGGC 1571
 DB 77027 TTTAAATAGTACAGTGAATGATTCCTTACAAATCTATTTGGATTTGGTGGTGGC 76968
 QY 1572 TTTAGTCTTACCGGCAATCACTAACCTTTTAAAGTATCCAAATACGACGAGGCGCA 1631
 DB 76967 TTAGATCTTACCGGCAATCACTAACCTTTTACCGATCCAAATACGACGAGGCGCG 76908
 QY 1632 ATGATGTGACCAATTAATCACTACAGCGGCTATGTCATTAATTTGAGAAACGAAGC 1691
 DB 76907 ATGATGTGACCAATTAATCACTACAGTGTATTAATTTACTATTAATGAGAAACGAAGT 76848
 QY 1692 AT---TGTCTACCTAATGSAATTAATTAATTAATTAATTAATTAATTAATTAAT 1748
 DB 76847 ATTAATCTCTCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 76788
 QY 1749 TACAAAGGTTGGTTGGCGAAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1808
 DB 76787 TACAAAGGTTGGTTGGCGAAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 76728
 QY 1809 TATAACCAACACAGAGATGTACTTCTACTTCACTTCACTTCACTTCACTTCACTTCACT 1868
 DB 76727 TATAACCAACACAGAGATGTACTTCTACTTCACTTCACTTCACTTCACTTCACTTCACT 76668
 QY 1869 GATAATTAACCAACAAAGTAACTATTTTTCAGCGGTACAGCAACCGGACCGCTAC 1928
 DB 76667 GATAATTAACCAACAAAGTAACTATTTTTCAGCGGTACAGCAACCGGACCGCTAC 76638
 QY 1929 AATCATTTAATTAAGCTTGGTCAAGAAATGGAAGTAAACCAAGGCGAAATTTGTGG 1988
 DB 76607 AATCATTTAATTAAGCTTGGTCAAGAAATGGAAGTAAACCAAGGCGAAATTTGTGG 76548
 QY 1989 GATCAAGATTAAGTCAACCGTACATTTAAAGTGAAGAACTTCCAAATTAAGGCGAAGT 2048
 DB 76547 GATTAAGATTAAGTCAACCGTACATTTAAAGTGAAGAACTTCCAAATTAAGGCGAAGT 76488
 QY 2049 GCGGTGGTTCTCGCAATGTTTCTCAATTAAGGAAATTTGACAGTCAAGCAATTAATCA 2108
 DB 76487 GCGGTGGTTCTCGCAATGTTTCTCAATTAAGGAAATTTGACAGTCAAGCAATTAATCA 76428
 QY 2109 AATGCCAATTTGTTGTTGTCGAATCAATCAATTAATTAAGGCGGTTCAATTTGG 2168
 DB 76427 AATGCCAATTTGTTGTTGTCGAATCAATCAATTAATTAAGGCGGTTCAATTTGG 76368
 QY 2169 ACAGGATTAAGCACTGTCAAAAGTGAATTTTACGATTAAGGAAAGTAAATTAATTAAT 2228
 DB 76367 ACAGGATTAAGCACTGTCAAAAGTGAATTTTACGATTAAGGAAAGTAAATTAATTAAT 76308
 QY 2229 CCAAAAACAAATTAAGTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2288
 DB 76307 CCAAAAACAAATTAAGTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 76248
 QY 2289 GGTTTAGCAAACTTAATGCAATGTCATTTAAGCAATCAAGGCAATTTCAATTAAGC 2348
 DB 76247 GGTTTAGCAAACTTAATGCAATGTCATTTAATGCAATCAAGGCAATTTCAATTAAGC 76188
 QY 2349 AACAAATCCACCAATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2408
 DB 76187 AACAAATCCACCAATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 76128
 QY 2409 AATCAAACTTGAACGTTATGTCATTTAAGCAATCAAGTCAATTTCTTTAAAAAC 2468
 DB 76127 AATCAAACTTGAACGTTATGTCATTTAAGCAATCAAGTCAATTTCTTTAAAAAC 76068
 QY 2469 AGCAATTTTTCGACCAAAATTCAGGAGACAAAGGCAACAGTGAACGTTGAAAATTCG 2528
 DB 76067 AGCAATTTTTCGACCAAAATTCAGGAGACAAAGGCAACAGTGAACGTTGAAAATTCG 76008
 QY 2529 ACTTGCACAAAGCTTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2588
 DB 76007 ACTTGCACAAAGCTTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75948
 QY 2589 ACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2645
 DB 75947 ACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75888
 QY 2646 GAGACGAAACAAAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2705
 DB 75887 GAGACGAAACAAAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75828
 QY 2706 AATTAAGTGGCAAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2765
 DB 75827 AATTAAGTGGCAAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75768
 QY 2766 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2825
 DB 75767 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75708
 QY 2826 AAGAAACCGGAAACCTTGAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2885
 DB 75707 AAGAAACCGGAAACCTTGAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75648
 QY 2886 TCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2945
 DB 75647 TCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75588
 QY 2946 AATTAAGTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3005
 DB 75587 AATTAAGTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75528
 QY 3006 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3065
 DB 75527 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75468
 QY 3066 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3125
 DB 75467 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75408
 QY 3126 GCGTTTCTGATACCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3185
 DB 75407 GCGTTTCTGATACCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75348
 QY 3186 GAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3245
 DB 75347 GAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75288
 QY 3246 GAGTGTGTTTCTGATACCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3293
 DB 75287 GAGTGTGTTTCTGATACCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75228
 QY 3294 GCGGCACTTGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3353
 DB 75227 GCGGCACTTGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75168
 QY 3354 GAAGCGGAAACAAAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3413
 DB 75167 GAAGCGGAAACAAAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75108
 QY 3414 TCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3473
 DB 75107 TCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 75048
 QY 3474 TTTGTAGTCAAGCAATTTGCGGTGTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3533
 DB 75047 TTTGTAGTCAAGCAATTTGCGGTGTAAGCAATTAATTAATTAATTAATTAATTAATTAAT 75008
 QY 3534 GATTTGATGCGTTCGTTGTTTCAAGGCAAGAAAGCAATTAATTAATTAATTAATTAATTAATTAAT 3593


```

Db 730 GGAATATATAGCGTTGATTTTGAGAGGAGGTCGCAATCCCGATCGGATGTTTACT 789
Qy 420 TATAGTTGTAAAGGAAATTAATTAAGATTAATTAATTAATTAATTAATTAATTA 479
Db 790 TATCAATTTGTAAAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 849
Qy 480 TACCATTAATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 539
Db 850 TATACATGCTGCTGTTGCAATTTGTCTGACGCTGACGCTGACGACGACGACGACGAC 909
Qy 540 AATATGATGAGAGTACTTATTCAGATTAAGCAAAATTAATTAATTAATTAATTAAT 599
Db 910 AATATGATGAGAGTACTTATTCAGATTAAGCAAAATTAATTAATTAATTAATTAAT 969
Qy 600 TCTGACGAGAGTCTTGGCGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 644
Db 970 AAGATGAAATGATGAGAGTCTGATGAAGCAAAAGCAAGCAAGCAAGCAAGCAAG 1029
Qy 645 GTTCCGCTGATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 704
Db 1030 CTTCCTATGCTTATCTGATGAGATTAAGCAAGTAAACATTAATTAATTAATTAAT 1089
Qy 705 AATGATATTCGATTTGAGAGCGATGTTCTAAAGCGGAGATTAATTAATTAATTA 764
Db 1090 AAGCGACGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1149
Qy 765 ATTCAGGCTCAAAAGGAGGAGATGATGCTGATGATTAATTAATTAATTAATTAAT 824
Db 1150 AAGGATGCTGCTGATGAGATTAAGTCTGATGATTAATTAATTAATTAATTAATTA 1209
Qy 825 AATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 884
Db 1210 AATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1269
Qy 885 TTTCAATGCTGCAAACTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 941
Db 1270 TTTCAATGCTGCAAACTTATTAATTAATTAATTAATTAATTAATTAATTAAT 1329
Qy 942 TCACTTACACCGGAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1001
Db 1330 ACATTTTAAAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1389
Qy 1002 GGGCTCTAATCTC-----AGATTCAGATTAATTAATTAATTAATTAATTAAT 1052
Db 1390 GGTACGCTTACCAAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1449
Qy 1053 GCAATATGAGTTCCTTGAAGAGAGATTAATTAATTAATTAATTAATTAATTAAT 1112
Db 1450 CAGTATTTAATGAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1509
Qy 1113 CCTATATTTATCTCAAGTTTAAAGATTAATTAATTAATTAATTAATTAATTAATTA 1172
Db 1510 GTAAATGCTTAAAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1569
Qy 1173 CAGGATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1232
Db 1570 AAGAGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1629
Qy 1233 GGTATTTTACGATTTCCAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1292
Db 1630 GGTATTTTACGATTTCCAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1689
Qy 1293 AGTAAATATGACCGTACTTGGAGATTAATTAATTAATTAATTAATTAATTAATTA 1352
Db 1690 GTTAAAGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1749
Qy 1353 ATTGTAAGAGAACTTCAAGCTCAAGCAAGGAGAAATTAATTAATTAATTAATTA 1412
Db 1750 ATTGTAAGAGAACTTCAAGCTCAAGCAAGGAGAAATTAATTAATTAATTAATTA 1809
Qy 1413 GGCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1472

```

```

Db 1810 GGTGACGCAAGATCATTTAGATCAACAGCCGATGAGACACCAAAACAAAGCTTT 1869
Qy 1473 AGTAAATTTGCTGCTTGGAGGAGAGGAGCTGTCATTTAAAGCAATTAATTAATTA 1532
Db 1870 AAGAGATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1929
Qy 1533 GATACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1592
Db 1930 GATCCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1989
Qy 1593 TTAACCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1652
Db 1990 TTAACCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2049
Qy 1653 ACTCAAGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1709
Db 2050 ACTCAAGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2109
Qy 1710 AATATATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1769
Db 2110 AATATATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2169
Qy 1770 ACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1829
Db 2170 ACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2229
Qy 1830 GGTACTTGTCTACTTCAAGCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 1889
Db 2230 GGTACTTGTCTACTTCAAGCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 2289
Qy 1890 AATCTATTTTCAAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1949
Db 2290 ACTTATGATTTTATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2343
Qy 1950 TCGAATTAAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2009
Db 2344 CCAATGAGCTTGGAGCTTCAAGCTGATTAATTAATTAATTAATTAATTAATTAATTA 2403
Qy 2010 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2069
Db 2404 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2463
Qy 2070 TCTCAATTAAGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2129
Db 2464 TCTCAATTAAGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2523
Qy 2130 CCAATTAAGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2189
Db 2524 CCAATTAAGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2583
Qy 2190 AAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2249
Db 2584 ACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2643
Qy 2250 TCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2309
Db 2644 TCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2703
Qy 2310 AATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2369
Db 2704 AATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2763
Qy 2370 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2429
Db 2764 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2823
Qy 2430 GTGATTTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2489
Db 2824 GTGATTTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2883
Qy 2490 CAGGAGCAAGGAGCAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2549
Db 2884 CAGGAGCAAGGAGCAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2943
```

QY 2556 ACTACATTCGAGATTTTAAGCTTAATAAGATGATGACGTTAAATTCAGCTTATTCQA 2609
 Db 2944 GCCACATTCGAGATTTTAAGCTTAATAAGATGATGACGTTAAATTCAGCTTATTCQA 3003
 QY 2610 GCTAGCTCAACATATGACCACT---CGCGCTTCATTAGACGAGAAACACGCCAACA 2666
 Db 3004 GCTAGCTCAACATATGACCACTGACCGCGCTTCATTAGACGAGAAACACGCCAACA 3063
 QY 2667 TCGGCAAGACATCGTTTCAACATTTGACGTAATGTAATGAGTGCGGCAAGGCA 2726
 Db 3064 TCGGCAAGACATCGTTTCAACATTTGACGTAATGTAATGAGTGCGGCAAGGCA 3123
 QY 2727 TTCCATTTACTTCATCTTATTTGGCTATTAAGCGATTAATTAATTTCCATGAC 2786
 Db 3124 TTCCATTTACTTCATCTTATTTGGCTATTAAGCGATTAATTAATTTCCATGAC 3183
 QY 2787 GCTGAGGCGATTAATATATCTGTTGCAACACAGGCAAGAACCCGAACCCCTTGAG 2846
 Db 3184 GCTGAGGCGATTAATATATCTGTTGCAACACAGGCAAGAACCCGAACCCCTTGAG 3243
 QY 2847 CAATTAATCTTGTGAAAGCAAGATATCAACCGTTATCAGATTAAGCTCAATTTACT 2906
 Db 3244 CAATTAATCTTGTGAAAGCAAGATATCAACCGTTATCAGATTAAGCTCAATTTACT 3303
 QY 2907 TTGGAATGACCAAGCTTATGAGGCTGCTTAAGCTTATTAATGAGTAATGATGAG 2966
 Db 3304 TTGGAATGACCAAGCTTATGAGGCTGCTTAAGCTTATTAATGAGTAATGATGAG 3363
 QY 2967 GAATTCGGCTTCATTAACCCATTAAGAGCAGAAATGACATGATTAATGATGAGCA 3026
 Db 3364 GAATTCGGCTTCATTAACCCATTAAGAGCAGAAATGACATGATTAATGATGAGCA 3423
 QY 3027 GAGCAAGCAGAACGAACTTGAAGCCAAACAGTTGACCGATCTGTAACCAACCA 3086
 Db 3424 GAGCAAGCAGAACGAACTTGAAGCCAAACAGTTGACCGATCTGTAACCAACCA 3483
 QY 3087 GGTGAGCCAAAGTGGCTGCAAGAGCAGG----- 3119
 Db 3484 AGTATGCAAGAGTGGCTGCAAGAGCAGGTTTCTGATACCTGCTGATCAAGG 3543
 QY 3120 ----- 3119
 Db 3544 CAGTTAGAGTATTAACAGCGAACAAGTTGAACGACCTGCTGATCAAGGATTAAG 3603
 QY 3120 -----AGAGCAGCTTCTGATACCTGCTGATCAAGGCTG 3158
 Db 3604 GCAAAAAAGTGGCTCAAAAAGAGCGGTGTTCTGATACCTGCTGATCAAGGCA 3663
 QY 3159 TTAAGCGCTTAGAGCCAAACAGCTGAACCTGCTGATCAACAAAAAGTAAAGCA 3218
 Db 3664 TTAAGCGCTTAGAGCCAAACAGCTGCTGATCAACAAAAAGTAAAGCA 3723
 QY 3219 AAAACAAAAAGTGGCTCAAAAAGAGAGTGTCTGATCCCTGCTGATCAAGG 3278
 Db 3724 AAAAAGTGGCTCAAAAAGAGCGGAGAGGTTTCTGATACCTGCTGATCAAGG 3783
 QY 3279 CTGTTCC---GCATTGAAGCCGACCTTAGGTTATGATGCCCAACGATATGAAAAA 3335
 Db 3784 CGGTTAAAGGTATTAAGAGTCAAACTGAGTTATTAATGCCCAACGATATGAAAAA 3843
 QY 3336 GATGCTGATGCTCAAGAGAGAGGAAAAACAGCAAAACAAAGAACTGATGACGCGT 3395
 Db 3844 GAACTCAAGATCA-----GAGAAACAGCAAAACAAAGAACTGATGACGCGT 3894
 QY 3396 TATTCAAATAGTGGCTATCAAGATTAATCTGCAACAGTAATAGTATGCTTTCTGTC 3455
 Db 3895 TATTCAAATAGTGGCTATCAAGATTAATCTGCAACAGTAATAGTATGCTTTCTGTC 3954
 QY 3456 GATGAATTAAGTATGCTTTTGTATGATCAAGACCAATGCGCGTGGGCAAAATATGCG 3515
 Db 3955 GATGAATTAAGTATGCTTTTGTATGATCAAGACCAATGCGCGTGGGCAAAATATGCG 4014

QY 3516 CAGATTAAGAGCAGTATGATGCTTCGCTGCTTATCAGACAGAGAAAAAGAAC 3575
 Db 4015 CAGATTAAGAGCAGTATGATGCTTCGCTGCTTATCAGACAGAGAAAAAGAAC 4071
 QY 3576 TTAAGCTCAATTTGGGCTGCAAAAAGCCTTAGCTAATGACAAATTTGGGCGATTTCTG 3635
 Db 4072 TTAAGCTCAATTTGGGCTGCAAAAAGCCTTAGCTAATGACAAATTTGGGCGATTTCTG 4131
 QY 3636 CATACCGCTTCATATATCTTGTATGAGACAGTTAAATAATCAAGGACATTAAGATG 3695
 Db 4132 CATACCGCTTCATATATCTTGTATGAGACAGTTAAATAATCAAGGACATTAAGATG 4191
 QY 3696 ATGTGCGGCTTTGCGCCATATCAATGCGGCGATTAACATTTGGTATGAAAGTGGAAAG 3755
 Db 4192 ATGTGCGGCTTTGCGCCATATCAATGCGGCGATTAACATTTGGTATGAAAGTGGAAAG 4251
 QY 3756 GGAATCAGTGGGATTAATTTGCTGAGAACAAAGCCGAAAAATTCATGCAAAAGGATA 3815
 Db 4252 GGAATCAGTGGGATTAATTTGCTGAGAACAAAGCCGAAAAATTCATGCAAAAGGATA 4311
 QY 3816 AATTAATGCGTGAATGCAATGATCACTGCTTATGAGGCAATTTGGGATTCAGCCTAT 3875
 Db 4312 AATTAATGCGTGAATGCAATGATCACTGCTTATGAGGCAATTTGGGATTCAGCCTAT 4371
 QY 3876 TTGAGATTAATGCTATTTTATGACGTAAGAAATTAATCAATCTGAGAGAGTGAAGTG 3935
 Db 4372 TTGAGATTAATGCTATTTTATGACGTAAGAAATTAATCAATCTGAGAGAGTGAAGTG 4431
 QY 3936 AAAAGCCTTACCTTCATTTATGCTATTAATGCTGATGATGATTAATTAATTA 3995
 Db 4432 AAAAGCCTTACCTTCATTTATGCTATTAATGCTGATGATGATTAATTAATTA 4491
 QY 3996 ACTCGAGATTAATTAATTAATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 4055
 Db 4492 ACTCGAGATTAATTAATTAATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 4551
 QY 4056 AACGCTAATGCTAATTAATTAATGCTTATGCTTATGCTTATGCTTATGCTTAT 4115
 Db 4552 AACGCTAATGCTAATTAATTAATGCTTATGCTTATGCTTATGCTTATGCTTAT 4611
 QY 4116 TGGCAAAAAGAGTGGGATTAAGGCAAGAACTTTTCAATTTCAAAATTTCCCTTTATC 4175
 Db 4612 TGGCAAAAAGAGTGGGATTAAGGCAAGAACTTTTCAATTTCAAAATTTCCCTTTATC 4671
 QY 4176 TCAAAATCTCAAGGCTCAACACTGCGCAACAGCAAAATGCGGCGTGAATTTGGCTAT 4235
 Db 4672 TCAAAATCTCAAGGCTCAACACTGCGCAACAGCAAAATGCGGCGTGAATTTGGCTAT 4731
 QY 4236 CGTGGTAAATTAATCAATTAATTTATGCTTATGATTAACAGATGCGTCAATCAGA 4295
 Db 4732 CGTGGTAAATTAATCAATTAATTTATGCTTATGATTAACAGATGCGGCGATC--- 4788
 QY 4296 TCCCACTTTTATTCATTAAT 4319
 Db 4789 ---CAACTTTTATTCATTAAT 4810

RESULT 6
 ADA09357
 ID ADA09357 standard; DNA, 4828 BP.
 XX
 XX ADA09357:
 DT 06-NOV-2003 (first entry)
 XX
 XX Haemophilus HAP protein gene from strain 1396B.
 XX
 XX ds; Haemophilus adherence and penetration protein; HAP; gene;
 KW antibacterial; immunostimulant; vaccine; Haemophilus influenza infection;
 XX influenza.
 XX
 XX Haemophilus influenzae; strain 1396B.
 XX

PN US2003073166-A1.
 XX 17-APR-2003.
 PD 22-FEB-2002; 200205-00080505.
 XX 25-AUG-1994; 94US-00296791.
 PR 20-APR-2001; 2001US-00839996.
 XX (GENE/) GENE J W S.
 PI Gene JWS;
 DR WPI: 2003-56738/53.
 DR P-PSDB: ADA09358.
 XX
 PT New recombinant Haemophilus adhesion and penetration protein, useful as a
 vaccine for prophylactic or therapeutic treatment of a Haemophilus
 influenzae infections, in generating antibodies for passive immunization.
 XX
 PS Claim 1; Fig 24; 14pp; English.
 CC The invention relates to a new recombinant Haemophilus adhesion acid
 CC penetration protein (HAP) appearing as ADA09348, ADA09350, ADA09352 and
 CC ADA09354 encoded by a nucleic acid appearing as ADA09349, ADA09351,
 CC ADA09353, ADA09355 and ADA09357. Also included are a recombinant nucleic
 CC acid that will hybridize under stringency conditions to the nucleic acid
 CC detailed above, an expression vector comprising transcriptional and
 CC translation regulatory nucleic acid operably linked to the HAP nucleic
 CC acid, a producing HAP by culturing a host cell transformed with the
 CC expression vector, a composition comprising a pharmaceutical carrier and
 CC HAP, inducing an immune response in a patient by administering the
 CC composition, a composition comprising an antibody capable of binding to
 CC HAP (or to a peptide selected from 35 peptide sequences appearing as
 CC ADA09353-ADA09353), and an antibody capable of binding to the peptides.
 CC The Haemophilus adhesion and penetration (HAP) protein is useful as a
 CC vaccine for prophylactic or therapeutic treatment of an H. influenzae
 CC infection (e.g. influenza), in generating an immune response in a
 CC patient, and in generating antibodies to the HAP protein. The antibodies
 CC may be used to diagnose the presence of H. influenzae infection in a
 CC sample or patient, in passive immunisation treatments, and to purify or
 CC separate HAP proteins or the H. influenzae organism from a sample. The
 CC present sequence is a HAP gene from a strain of H. influenzae.
 XX
 SQ Sequence 4828 BP; 1641 A; 896 C; 993 G; 1298 T; 0 U; 0 Other;
 Query Match 68.8%; Score 2970; DB 8; Length 4828;
 Best Local Similarity 82.3%; Freq. No. 0;
 Matches 3581; Conservative 0; Mismatches 645; Indels 124; Gaps 10;
 QY 1 TCATATGCTGTTACTA-GTATTTTAAACGAAATTAATTAATTAATTAACAT 59
 DB 253 TCATATGCTGTTAAACGATTTTAAACGAAATTAATTAATTAATTAACAT 312
 QY 60 ATGAAAAAACTGATTTGCTTAATTTTAAACGCTGATTCATTAGGATAGTA 119
 DB 313 ATGAAAAAACTGATTTGCTTAATTTTAAACGCTGATTCATTAGGATAGTA 372
 QY 120 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 179
 DB 373 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 432
 QY 180 GCCGAGAAATTAAGGAGTTCAGATTGGGAGCTCAAAATTAATTAATTAATTA 239
 DB 433 GCCGAGAAATTAAGGAGTTCAGATTGGGAGCTCAAAATTAATTAATTAAT 492
 QY 240 GGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 299
 DB 493 GGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 552
 QY 300 GGTAAAGGAGTTCAGATTGGGAGCTCAAAATTAATTAATTAATTAATTAAT 359
 DB 553 GGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 612

QY 360 GATATTAACGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 419
 DB 613 GGTATTAACGATTTGATTTGGTGCAGAGGAAACACCCGATCAATCGTTTACT 672
 QY 420 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 479
 DB 673 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 732
 QY 480 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 539
 DB 733 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 792
 QY 540 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 599
 DB 793 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 852
 QY 600 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 659
 DB 853 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 912
 QY 660 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 719
 DB 913 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 972
 QY 720 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 779
 DB 973 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1032
 QY 780 GGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 839
 DB 1033 GGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1092
 QY 840 GGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 899
 DB 1093 GGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1152
 QY 903 AATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 959
 DB 1153 AATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1212
 QY 960 GGTAAAGGAGTTCAGATTGGGAGCTCAAAATTAATTAATTAATTAATTAAT 1019
 DB 1213 GGTAAAGGAGTTCAGATTGGGAGCTCAAAATTAATTAATTAATTAATTAAT 1272
 QY 1020 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1079
 DB 1273 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1332
 QY 1080 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1139
 DB 1333 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1386
 QY 1140 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1195
 DB 1387 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1446
 QY 1200 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1259
 DB 1447 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1506
 QY 1260 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1319
 DB 1507 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1566
 QY 1320 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1379
 DB 1567 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1626
 QY 1380 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1439
 DB 1627 TCGAAGCGTGGCTGTCACATTTTGGATGATGACCAATTAATTAATTAAT 1686

[illegible]

Db	2749	GAATATGGCACTTGGCAATGCTCTGAGGATGACTACATTGGCAATTTAAACCTTAATAT	2808
Oy	2580	AGTACGATCAAGTTAAATTCAGCTTATTCAGCTACTCAACAAATGCGCACG---CCG	2636
Db	2809	AGTACGTTTACGTAAATTCAGCTTATTCAGCTACTCAACAAATGCGCACGCGC	2868
Oy	2637	CGTTCATTAGAGCGGAAACACCGCAACATCGGGCAACATCGTTGCACACATCGACA	2696
Db	2869	CGTTCATTAGAGCGGAAACACCGCAACATCGGGCAACATCGTTGCACACATCGACA	2928
Oy	2637	GTAAATGTAATTTAGTGGGCAAGGCAATTCGAATTACTTCATCTTATTTGGCTAT	2756
Db	2929	GTAAATGTAATTTAGTGGGCAAGGCAATTCGAATTACTTCATCTTATTTGGCTAT	2988
Oy	2757	AAAGGCAATTAATTAATTAATTCACATGACGCGGAGGCGCATTTACCAATTAATCGTGGC	2816
Db	2969	AAAGGCAATTAATTAATTAATTCACATGACGCGGAGGCGCATTTACCAATTAATCGTGGC	3048
Oy	2817	AACACAGCAAGAACCCGCAACCCCTTGACCAATTACCTTGGTTAAAGCAAGATAT	2876
Db	3049	GACACAGCAAGAACCCCTGACCCCTTGACCAATTAATTAAATGAGGCTGATAT	3138
Oy	2877	CACCGCTTATACATTAAGTCAAATTTACTCTTGAAAGACACGCTTGATGACGTGCA	2936
Db	3109	CACCGCTTGCATGATAGTAAAGTAAATTTACTCTTGAAAGACACGCTTGATGACGTGCA	3168
Oy	2937	TTACGTTAAATTAATGTGAGAGATGATGCGCATTCGCTTGCAATACCAATTAAGAG	2996
Db	3169	TGGCGTTATGATTTAGTGAAGAAAGACGGGAAATTCGCTTGCAATACCAATTAAGAG	3228
Oy	2997	CAGAAATTCACAAATGATTTAGTGAAGACGACGACGACGACGACATTTGAAGCCAA	3056
Db	3229	CAGAAATTCGCCATGATTTAGTGAAGACGACGACGACGACGACATTTGAAGCCAA	3288
Oy	3057	CAAGTTGAACCGATCTGTAAACCAACACGCGGACCAAAAGTGGCGTCAAGAGCA	3116
Db	3289	CAAGCTGAATCGATCTGTAAACCAACACGCGGACCAAAAGTGGCGTCAAGAGCA	3342
Oy	3117	GGCAGGAGCGGTTCCGTACCTACCTCGCTGATCAACCGCTGTAAACGATTAAGAGCC	3176
Db	3343	---AGAGGGGGTTTCTGATCCCGCTGATCAACCGCTGTAAACGATTAAGAGCC	3399
Oy	3177	AAACAAAGCTGAATCTGATCTGTAAACCA---AAAGTAAAGCGAAACAAAAAG	3231
Db	3400	GAACGTGAACGATTAATGCCCCAACAGAAAGTGGACAAAGGCGTGAATATCAAAAGTA	3459
Oy	3232	TGCGGTCAAAAGAGCA---GTGTTTTCGATCCCGCTGATCAAAAGCGTCTCGCA	3287
Db	3460	ACTGCACTTAAACCAAAAGACGACGACGATTTAAACACTGACATTAAGCAAAATTAGTC	3519
Oy	3288	TTAAGAGCGGCACTTGAGGTTATTTGATCCCGAC-----	3321
Db	3520	TTGGCAATCTGATTTGGTGAAGAAAGAAACCGCTCAGATTTGATTTGCTAATGCAAAATTA	3579
Oy	3322	-----AGCAATCGGAA	3333
Db	3580	CTTCACTTGAATTTAACACAACTAGAAAAACCTTGACATGCTGAGACAGACGAA	3633
Oy	3333	AAAGATCGCTAGCTCAAGAGAAAGACGAGAAACAAAGCAAAACAAAGACTGATCAG	3392
Db	3640	AAAGATCGTAAAGCTCAAGAGCAAGCAAAAGCAACGCAAAACAAAAACTGATCAG	3699
Oy	3393	CGTTATTCAAATAGTGGTATTCAGAAATTATTCGCAACAGTAAATGATATGCTTTCGT	3455
Db	3700	CGTTATTCAAATAGTGGTATTCAGAAATTATTCGCAACAGTAAATGATATGCTTTCGT	3755
Oy	3453	CAAGATGAATTAAGATCGCTTTTGTGTGATCAAGACCAATCTCGCTGGACAAATATC	3512
Db	3760	CAAGATGAATTAAGATCGCTTTTGTGTGATCAAGACCAATCTCGCTGGACAAATATC	3818
Oy	3513	GCACAGGATAAAGACGCTATGATTTGATGCGCTTCGTGTTATAGACGCAAAAGCG	3572

```

Db 3820 TCACAGGATTAAGACGTTATGATTCGATCGCTTCCTGTTT---CAGCAGAAAACG 3876
Qy 3573 AACTACGTCMAATGCGGATGCAAAAAGCCTTAGCTAATGAGCAAAATGGGCGAGTTTC 3632
Db 3877 AACTGCTCAATGCGGATGCAAAAAGCCTTAGCTAATGAGCAAAATGGGCGAGTTTC 3936
Qy 3633 TCGCATAGCGCTTCAGATTAATCCTTTGATGAACAGGTTAAAATCAACCGCATTAACG 3692
Db 3937 TCGCATAGCGCTTCAGATTAATCCTTTGATGAACAGGTTAAAATCAACCGCATTAACG 3996
Qy 3693 ATGATGCGGGTTTGGCCATATCATGCGGCGATTTACAGATTTGGTGTAAAGTGGGA 3752
Db 3997 ATGATGCGGGTTTGGCCATATCATGCGGCGATTTACAGATTTGGTGTAAAGTGGGA 4056
Qy 3753 ACGGGAATCAGTGCAGATTAATGCTGAAGAACCAACCGCAAAATTCATCGAAAACG 3812
Db 4057 ACGGGAATTAAGTGCAGATTAATGCTGAAGAACCAACCGCAAAATTCATCGAAAACG 4116
Qy 3813 ATAAATTAATGCGGATGCAATGCAATGATTCAGTTCCGTTAAGGCAATTCAGCT 3872
Db 4117 ATAAATTAATGCGGATGCAATGCAATGATTCAGTTCCGTTAAGGCAATTCAGCT 4176
Qy 3873 TATTTGAGATTATATGCTATTTATTTAGAGCGTAAATATCATCTGAGAGAGTAGA 3932
Db 4177 TATTTGAGATTATATGCTATTTATTTAGAGCGTAAATATCATCTGAGAGAGTAGA 4236
Qy 3933 GTGAAAACGCTAGCTGCTGATTTAATGCTAATATGCTGCAATTCAGTTGATTAACA 3992
Db 4237 GTGAAAACGCTAGCTGCTGATTTAATGCTAATATGCTGCAATTCAGTTGATTAACA 4296
Qy 3993 TTTTCTCCAGCATATATGAGGTTAAAGCTTTTCTTCTGCTAATATTTGATGTT 4052
Db 4297 TTTTCTCCAGCATATATGAGGTTAAAGCTTTTCTTCTGCTAATATTTGATGTT 4356
Qy 4053 TCAAGCGCTAGCTAATGCAACGAGTAAATTCAGAGTGTGCAACACATTTGAGAGT 4112
Db 4357 TCAAGCGCTAGCTAATGCAACGAGTAAATTCAGAGTGTGCAACACATTTGAGAGT 4416
Qy 4113 TATTTGCAAAAAGAGTGGATTAAGGCAAAATTTTACATTCGAAATTCGCTTTT 4172
Db 4417 TATTTGCAAAAAGAGTGGATTAAGGCAAAATTTTACATTCGAAATTCGCTTTT 4476
Qy 4173 ATTCGAAATTCAGAGTTCGCAATCTGCGAATGCAAGAAATGCGGCTGAAATTCGGC 4232
Db 4477 ATTCGAAATTCAGAGTTCGCAATCTGCGAATGCAAGAAATGCGGCTGAAATTCGGC 4536
Qy 4233 TATGCTGTAAATTAATCAATATTTTAT 4262
Db 4537 TATGCTGTAAATTAATCAATATTTTAT 4566

```

```

RESULT 7
ADA09355
ID ADA09355 standard; DNA; 4822 BP.
AC ADA09355;
DT 06-NOV-2003 (first entry)
DE Haemophilus HAP protein gene from strain 3219B.

```

```

XX db; Haemophilus adherence and penetration protein; HAP; gene;
XX Km antibacterial; immunostimulant; vaccine; Haemophilus influenza infection;
XX influenza.
XX OS Haemophilus influenzae; strain 3219B.
XX PN US003073166-A1.
XX PD 17-APR-2003.
XX PF 22-FEB-2002; 2002US-00080505.
XX

```

```

PR 25-AUG-1994; 94US-00236791.
PR 20-APR-2001; 2001US-00839996.
PA (GENE/) GENE J M S.
PI Gene JMS;
XX MPI; 2003-567308/53.
XX P-PSDB; ADA09356.
DR
DR
PT New recombinant Haemophilus adherence and penetration protein, useful as a
PT vaccine for prophylactic or therapeutic treatment of a Haemophilus
PT influenzae infections, in generating antibodies for passive immunization.
PS
PS Claim 1; Fig 22; 149pp; Eng.-lit.
CC The invention relates to a new recombinant Haemophilus adhesion and
CC penetration protein (HAP) appearing as ADA09348, ADA09350, ADA09352 and
CC ADA09354 encoded by a nucleic acid appearing as ADA09349, ADA09351,
CC ADA09353, ADA09355 and ADA09357. Also included are a recombinant nucleic
CC acid that will hybridize under stringency conditions to the nucleic acid
CC detailed above, an expression vector comprising transcriptional and
CC translation regulatory nucleic acid operably linked to the HAP nucleic
CC acid, producing HAP by culturing a host cell transformed with the
CC expression vector, a composition comprising a pharmaceutical carrier and
CC HAP, inducing an immune response in a patient by administering the
CC composition, a composition comprising an antibody capable of binding to
CC HAP (or to a peptide selected from 35 peptide sequences appearing as
CC ADA09359-ADA09393), and an antibody capable of binding to the peptides.
CC The Haemophilus adhesion and penetration (HAP) protein is useful as a
CC vaccine for prophylactic or therapeutic treatment of an H. influenzae
CC infection (e.g., influenza), in generating an immune response in a
CC patient, and in generating antibodies to the HAP protein. The antibodies
CC may be used to diagnose the presence of H. influenzae infection in a
CC sample or patient, in passive immunisation treatments, and to purify or
CC separate HAP proteins or the H. influenzae organism from a sample. The
CC present sequence is a HAP gene from a strain of H. influenzae.
XX
XX Sequence 4822 BP; 1652 A; 818 C; 982 G; 1370 T; 0 U; 0 Other.
SQ
Query Match 54.1%; Score 2770.2; DB 8; Length 4822;
Best Local Similarity 73.6%; Pred. No. 0;
Matches 3466; Conservative 0; Mismatches 813; Indels 76; Gaps 13;
Qy 1 TCAATAGCTGTTAATA GATTTTATATGCAAAATTTCTTAATTAATTAACATT 59
Db 328 TCAATAGCTGTTAATA GATTTTATATGCAAAATTTCTTAATTAATTAACATT 387
Qy 60 ATGAAAAAACTGTAATTCGCTTAATTTTAAACGCTGATTTGATGGATAGTA 119
Db 388 ATGAAAAAACTGTAATTCGCTTAATTTTAAACGCTGATTTGATGGATAGTA 447
Qy 120 TCGCAAGCGTGGGCTGCTACACTATTTTGGATGATTAACAATAATATGATTTT 179
Db 448 TCGCAAGCGTGGGCTGCTACACTATTTTGGATGATTAACAATAATATGATTTT 507
Qy 180 GCCGAGATTAAGGAAAGTTTACAGTGTGGGCTGAAATTTAAGCTTTATACAAACA 239
Db 508 GCCGAGATTAAGGAAAGTTTACAGTGTGGGCTGAAATTTAAGCTTTATACAAACA 567
Qy 240 GGGCAATTAAGTGGCAATGCAAAAGCCCGATGATTTTGTGATGCTA 299
Db 568 GGGCAATTAAGTGGCAATGCAAAAGCCCGATGATTTTGTGATGCTA 627
Qy 300 CGTACGCGGTGGAGCGCTTGTTGAAATCATATATTTGAGCGTGGCATTAACGTA 359
Db 628 CGTACGCGGTGGAGCGCTTGTTGAAATCATATATTTAGCGTGGCATTAACGTA 687
Qy 360 GATATACAGATGTTTATTTTGGTGCAGGGAAACACCCGATTCACATCGTTTACT 419
Db 688 GATATACAGATGTTTATTTTGGTGCAGGGAAACACCCGATTCACATCGTTTAAAG 747
Qy 420 TATTAAGTGTAAACGAATAATACAAAAGATATTTACATCTTATGAGAGCAT 473

```


1596 ACCTTTAAAGCTATCCAAATATACGAGCGAGGCGCATGATTGTGACCAATATACACT 1655
1976 ACCTTTAAAGCTATCCAAATATACGAGCGAGGCGCATGATTGTGACCAAGCGCTCT 2035
1656 CAAGCCGCTATATGCTATTTATCTGGAGACGAAGCATTTGTTCTACTATATGAAATAT 1715
2036 CAAGCCGCTATATGCTATTTATCTGGAGACGAAGCATTTATCA-----GAT 2083
1716 ATTAATTAATCTGATTAAGAAAGAAATGCTCTCAAGGTTGTTGGCGAAACGAT 1775
2084 AGCAACACTTACTATTAATAAAGATATGCAATTAAGCGTGGTTGTGTGACAGAT 2143
1776 AAAATTAACACAAATGCGCATTAACCTTATTTAAACCAACCAAGAGATGCTACT 1835
2144 AAAGCTAAACAAATGCTGTGTTTAAATGGAATATTAACAGATTAATGCAAAATCAT 2203
1836 TTGCTACTTCAAGTGTGTAATTAATTAAGGCAATTAACCAACCAAGGTAACCTA 1895
2204 TTGTTGCTTCTGAGGAGCAAAATTAACGCAATATCAAGCAAAATGTTGTGCTTA 2263
1896 TTTTTCAGGCTAGACCGACACCGACGCTCAATCATTTAAATTAAGTGTGCAAGAA 1955
2264 GTTTTATGTGTGTGTCACACCTCTATGCTTCAATCATTTAAGAAAGACCTTGTCTAC 2323
1956 ATGAGAGTATACCAACAGGCAAAATGTTGTGAGATCAAGATTGATCAACCGTACATT 2015
2324 ATGAGAGTATCCCAACAGGCAAAATGTTGTGAGATCAAGATTGATCAACCGTACATT 2383
2016 AAAGCTAAACCTTCAATTAATTAAGGCGAAATGCGGTTGTTCTGCAATGTTCTTCA 2075
2384 AAAGCTAAACCTTCAATTAATTAAGGCGAAATGCGGTTGTTCTGCAATGTTCTTCA 2443
2076 ATGAGAGGAAATTTGACAGTCAAGCAATATGCAATTAATGCAATTTGTTGTGCAAT 2135
2444 ATGAGAGGAAATTTGACAGTCAAGCAATATGCAATTAATGCAATTTGTTGTGCAAT 2503
2136 CAACAAATATCAATTTGACAGCTTCAAGTGTGCAATTAATGCAATTTGTTGTGCAAT 2195
2504 CAACAAATATCAATTTGACAGCTTCAAGTGTGCAATTAATGCAATTTGTTGTGCAAT 2563
2196 GATTTAACGATACAAATGTTATTTATTTATTAATCAAAACAAATCAATCAATGCTTAT 2255
2564 GATTTAACGATACAAATGTTATTTATTTATTAATCAAAACAAATCAATCAATGCTTAT 2623
2256 AATTTAATGATTAATGCAACGCGATTAATTAAGTTAGCAAACTTAATGCAATGTC 2315
2624 AATTTAATGATTAATGCAACGCGATTAATTAATGCAAACTTAATGCAATGTC 2683
2316 AATTTAATGATTAATGCAACGCGATTAATTAATGCAAACTTAATGCAATGTC 2375
2684 AATTTAATGATTAATGCAACGCGATTAATTAATGCAAACTTAATGCAATGTC 2743
2376 GCACTTTCGACATTAATGCAACGCGATTAATTAATGCAAACTTAATGCAATGTC 2435
2744 AACTTTCAATCAAGCAAAATGCAACGCGATTAATTAATGCAAACTTAATGCAAT 2803
2436 TTAACGATTAATGCAACGCGATTAATTAATGCAAACTTAATGCAAACTTAATGCAAT 2495
2804 TTAACGATTAATGCAACGCGATTAATTAATGCAAACTTAATGCAAACTTAATGCAAT 2863
2496 GCAACAGCAACAGGATGTTGGAATGCGATTTGCAATGCTTGAATGCTTGAATGCT 2555
2864 GCAACAGCAACAGGATGTTGGAATGCGATTTGCAATGCTTGAATGCTTGAATGCT 2923
2556 TTGCAAGATTAATGCAACGCGATTAATTAATGCAAACTTAATGCAAACTTAATGCAAT 2615
2924 TTGCAAGATTAATGCAACGCGATTAATTAATGCAAACTTAATGCAAACTTAATGCAAT 2983
2616 TCAACATTAATGCAACGCGATTAATTAATGCAAACTTAATGCAAACTTAATGCAAT 2669
2984 TCAACATTAATGCAACGCGATTAATTAATGCAAACTTAATGCAAACTTAATGCAAT 3043
2670 GCAACATGCTTTCAACATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 2729

3044 GCAGAACATGCTTTCAACATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3103
2730 CAATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 2789
3104 CAATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3163
2790 GAGGCGATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 2849
3164 GAGGCGATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3223
2850 TTAACCTTGTGGAACCAATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 2909
3224 TTAACCTTGTGGAACCAATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3283
2910 GAAATGCAACGTTGATGCAAGTGTGATTAATGCAAACTTAATGCAAACTTAATGCAAT 2969
3284 GAAATGCAACGTTGATGCAAGTGTGATTAATGCAAACTTAATGCAAACTTAATGCAAT 3343
2970 TTCCGCTGCAATTAACCAATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3029
3344 TTCCGCTGCAATTAACCAATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3403
3330 CAAGCAAGCAATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3089
3404 CAAGCAAGCAATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3463
3090 GAGGCGATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3149
3464 GAGGCGATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3514
3150 CAAGCTGTTTAAAGCAATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3209
3515 CAAGCTGTTTAAAGCAATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3571
3210 AGTAAGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3269
3572 AGTAAGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3631
3270 GATCAAGCTGTTGCAATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3329
3632 GATCAAGCTGTTGCAATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3682
3330 GAAAGCAATGCTTACGCTCA---GAAAGCAATGCTTACGCTCA---GAAAGCAATGCTTACGCTCA 3386
3683 GAAAGCAATGCTTACGCTCA---GAAAGCAATGCTTACGCTCA---GAAAGCAATGCTTACGCTCA 3742
3387 ATGAGCGCTTATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3446
3743 ATGAGCGCTTATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3802
3447 TCTGCTCAAGATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3506
3803 TCTGCTCAAGATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3862
3507 AATATGCAAGATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3566
3863 AATATGCAAGATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3923
3567 AATATGCAAGATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3626
3920 AATATGCAAGATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3979
3627 GTTTCGCTGATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3686
3980 GTTTCGCTGATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 4039
3687 TTAACATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3746
4040 TTAACATTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 4099
3747 GTGGAAGCGGAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAACTTAATGCAAT 3806


```

Db 4100 GTGGGTGGGGAATTAGTCGAGTAATAATGCTGAGAAACAAAGCCGAAATTCATCGA 4159
Qy 3807 AAAGCATTAATTAATGCGTAATGCAAGTATCAATTCGTTAGGCGCATTTGGCATT 3866
Db 4160 AAAGCATTAATTAATGCGTAATGCAAGTATCAATTCGTTAGGCGCATTTGGCATT 4219
Qy 3867 CAGCCTTAATTTGAGAGTTAATCGTAATTTATGACGTAATAATTAATCAATCGAGAA 3926
Db 4220 CAGCCTTAATTTGAGAGTTAATCGTAATTTATGACGTAATAATTAATCAATCGAGAA 4279
Qy 3927 GTAGAGTGAATAAGCCTACCTGCTGATTAATGCTTAATATGCTGATTAATGCTGAT 3986
Db 4280 GTAGAGTGAATAAGCCTACCTGCTGATTAATGCTTAATATGCTGATTAATGCTGAT 4339
Qy 3987 TATACATTTATCCGACAGATTAATATGAGCTTAATTCCTGCTGATTAATGCTGAT 4046
Db 4340 TATACATTTATCCGACAGATTAATATGAGCTTAATTCCTGCTGATTAATGCTGAT 4399
Qy 4047 GATGTTTCAAAAGCTTAACGTACAAACAGGTAATTCACGCTGCTGACAAACCATTT 4106
Db 4400 GATGTTTCAAAAGCTTAACGTACAAACAGGTAATTCACGCTGCTGACAAACCATTT 4459
Qy 4107 GAGCTTAATGCGCAAAAGAGTGGATTAAGGCAAAATTTTACATTTCCAAATTCCTC 4166
Db 4460 GAGCTTAATGCGCAAAAGAGTGGATTAAGGCAAAATTTTACATTTCCAAATTCCTC 4519
Qy 4167 GCTTTATCTCAAAATCTCAAGGTTCAACACTCGGCAACAGCAAAATGTTGGGCGTGA 4226
Db 4520 GCTTTATCTCAAAATCTCAAGGTTCAACACTCGGCAACAGCAAAATGTTGGGCGTGA 4579
Qy 4227 TTGGGCTATCGTTGGTAAATATCAAC 4252
Db 4580 TTGGGCTATCGTTGGTAAATATCAAC 4605

```

RESULT 9

ACAA1657
ID ACAA1657 standard; DNA; 4350 BP.

AC ACAA1657;

XX 19-JUN-2003 (first entry)

DE Prokaryotic essential gene #23314.

KM Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

XX Neisseria meningitidis.

PN W0200277183-A2.

XX 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00672851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KJ, Zyskind JW,

PI Wall D, Tirwick JD, Carr GJ, Yamamoto K, Forsyth RA, Xu HJ.

DR WPI; 2003-029926/02.

DR P-PDSB; ABU37787.

PT New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 29527; 1766bp; English.

PS The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from MPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 4350 BP; 1237 A; 1186 C; 1072 G; 855 T; 0 U; 0 Other;

Query Match 31.3%; Score 1353.8; DB 7; Length 4350;
 Best Local Similarity 59.7%; Pred. No. 0;
 Matches 2586; Conservative 0; Mismatches 1527; Indels 216; Gaps 10;

```

Qy 81 CTTAATTTTAAACCGCTGATTTCAATAGGATGATGCGCAACGCTGGCTGATC 140
Db 73 CCTGCTTACTTACCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 132
Qy 141 ACTTAATTTTGGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 200
Db 133 ACTTAATTTTGGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 192
Qy 201 ACGTTGGGCTCAAAATTAAGTTTATTAACAAACAGGCAATTAATTTGGACATCA 260
Db 193 GCAATGGGCGGAAAGATTAATTAAGTTTATTAACAAACAGGCAATTAATTTGGACATCA 252
Qy 261 ATGACAAAGCCCGATGATTAATTTTCTGATGATGATGATGATGATGATGATGATG 320
Db 253 ATGACAAAGCCCGATGATTAATTTTCTGATGATGATGATGATGATGATGATGATG 312
Qy 321 GTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 380
Db 313 GTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 372
Qy 381 GGTGCAAGGAAACCAACCCGATCAATCGTTTACTTAATTAATTAATTAATTAAT 440
Db 373 GGTGCAAGGAAAGAAATCCGATCAGACCGTTTCTTAACAAATTAATTAATTAAT 432
Qy 441 AACTCAAAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 500
Db 433 AATTAAGCTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 492
Qy 501 AATTCGTTACAGAGGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 560
Db 493 AATTTTCAAGATGACAGACCTGTGAAATTAAGAGTGAAGAGGGAATTAATTAAT 552

```

QY 561 TCAGATAGACAAATAATCCAGAACGTTGCGATCGGCTTCGAGCGGAGTTTGGCGA 620
 DB 553 TCGATTAAGAAAAATATCCGAGCGGTGCCATCGGCTCGAGACACCATATG66GT 612
 QY 621 AATGATAGACAAAGGCGACCAAGTTCCGCGATATCATATCTGACAGCTGCAAT 680
 DB 613 TATGATGATGACAAACGCGCATTTATCTCTCCGCGCATGCTTAATTTGGGCGCAAT 672
 QY 681 ACACAGATCAGCGTGGAGCAGGTAAATGAAATGTAATTTGGAGCGCAATGCTGTA 740
 DB 673 ACACATTTGAGGGGTTGGGAAATATATGAGTAAATGTAATTTGGAGCGCGATGCGCAT 732
 QY 741 GCGGAGAAATATGCTCCATTCAGATGCGAGCTCAAGAGGCGGACAGTGTCTCCAGT 800
 DB 733 GCGACGACATATGCGCTTATGCGATTCGAGGTGGGCGAGCGACAGCGGTTCCCATG 792
 QY 801 TTTATTTATGATGCTGAAAAAAGAAAAATGTTATTTAAATGGAATTTACGGGAAAGGAC 860
 DB 793 TTTATTTATGACAAACAAACAAATTAATGCGTGTCAACGAGATTTTAAACCGGCTAC 852
 QY 861 CCTTTGAGGCGAAAGAAAAATGGTTCAATTTGTTGCGCAATCTTAATTT---TGATGAA 917
 DB 853 CCTATTCGCGGAGGAAACGAGTTTCAGCTGATAGCAAGATTTGTTTACGATGAC 912
 QY 918 ATTTTGAAGAGATTTACATACATCACTTTACACCGAGCTGATATGAGTGTAGACA 977
 DB 913 ATTTTGAAGAGGAGATACACATACCGCTTTTGAACCGGCGAGTAAACGACATTTTCC 972
 QY 978 ATTTGTAAGATGATATGATGATGAGGAGTCTATATCTCAAGAAATCAAGAAATACCATCGAA 1037
 DB 973 TTTACATCCAAACAAACAGGTAGCGGAGTACGAGTAAACAAACAAAGGTTTCCAT 1032
 QY 1038 ATTTAAATTTACGTTAGCAATATGATTT-----ACCTTTGAAGAGAGGATTA 1088
 DB 1033 CCAAGCTTTAAAGTACAGACAGTCCGACTGTTGACGAATTTGATTAATCTGATTA 1092
 QY 1089 GTTCATATCTGATATGATGAGACGACCTAATTTATTTCTCCAGTTTAAACATATGAA 1148
 DB 1093 GAACCATTTTACGCGGCGAGGGGGTGTAAATCAGTACCGTCAAGGTTAAACAGGTTAA 1152
 QY 1149 ACGCTATTTTATGATTCACAAACAAAGATCATTAATTTTGGAGTACATCTTACCA 1208
 DB 1153 AACCTTCTTTATGATTTACGCGACGCAACATCATTTATCAACAAACATCAACAA 1212
 QY 1209 GGGCGGAGTGTCTTTATTTTGAAGGTAATTTTACATGATCTCAAACTCTTACCAACT 1268
 DB 1213 GGGCGGAGGAGTGTATTTATTTGAAGGTAATTTTACGCTCGCTGCTGAAACAAAGAACG 1272
 QY 1269 TGGCAGAGAGCTGGCATACATGTAAGTAAATAGCACGCTTACTTGAAGTAAATGAC 1328
 DB 1273 TGGCAGAGCGCGGCGTTCATATCAGTAAAGACATGACCTTACTTGAAGTAAAGCGC 1332
 QY 1329 GTGGAATGATGACCTTTCTAAATTTGTAAGAAAGAAAGATGACGCTTAAAGCGAAAGG 1388
 DB 1333 GTGGCAACGACCGCTTCCAAATTTGGAAAGGACAGCTGACGCTTAAACCGAAAGG 1392
 QY 1389 GAAATTAAGTGTCTGATCAGCGTAGCGATGTAAGTCAATTTTGAAGACAGACGAC 1448
 DB 1393 GAAACCAAGGCTGATGAGGTGGGAGAGGTACATCTTTTGAATGAGAGGCGAC 1452
 QY 1449 GATTAAGGCAACAAACAGCTTTGTAAGAAATGAGCTTGTGTAAGCGGCGAGAGGAGTGT 1508
 DB 1453 GATTAAGGCAACAAACAGCTTTGTAAGAAATGAGCTTGTGTAAGCGGCGAGGAGTGT 1512
 QY 1509 CAATTAAGATGATTAACATTTGTAAGAAATGTAATTTTGGCTTGTGAGTGT 1568
 DB 1513 CAATGATGCGATTAATGATTAACCGGCAACACTCTATTTGCGCTTGTGAGGCGGA 1572
 QY 1569 CGTTAGATCTTAAAGCGGCAATTTACCTTAAACCTTAAACGTAATCCAAATAGCGAGCGG 1628
 DB 1573 CGTTGATTTAAAGCGGCAATTTGCTTGTGTAAGCTTAAATTAATCCGATGAAGG 1632
 QY 1629 GCAATGATTTGTAACATTAATGACATCAAGCGGCTAAATGCTATTTCTGGGAAAGAA 1688

DB 1633 GCGATGATTTGCAATCAATATCCCAACAAACATCCACCGCTTACATTAACGAGATGA 1692
 QY 1689 AGATTTGTTCTACTATATGAAATTAATTTATTAATCTGATTAACGAAAGAAATTTGC 1748
 DB 1693 AGATTTTACAAACGAGTGTATGAATTAATTAATGATTTATTAATCAAGAAATTTGC 1752
 QY 1749 TACACGAGTGTGTTGGCGAAACAGATTAATTAATTAACAAATGCGCATTTAACTTAT 1808
 DB 1753 TACACGAGTGTGTTGGCGAAAGATTAACAAACAAACGAGGCGGCTTCACTTGT 1812
 QY 1809 TATTAACCAACACAGAGATGTAATTTGCTACTTTGAGTGTGTAACAAATTTAAAGGC 1868
 DB 1813 TACAGCGCGCGGAGAGACGCGACCGCTGCTTTCCGCGGAAACAAATTTAAACGCG 1872
 QY 1869 GATATTTCCAAACAAAGATTAATTTTTCAGGAGTAAACGACCGGACGCTAC 1928
 DB 1873 AACATTCAGCAACAAACGCAACCTGTTTTCAGGCGAGCGACCGCGCTTAC 1932
 QY 1929 AATCATTAATTAACGTTGTGTAAGATGTAATTCACAGAGCGCAATTTGTGTG 1988
 DB 1933 AATCATTTAGAGAGCGGAGTGTGTAAGATGTAATTCACAGAGCGCAATTTGTGTG 1992
 QY 1989 GATCAGATGATGATCAACGCTTAACTTTAAAGCTGAAACCTTCAATTTAAAGCGGAGT 2048
 DB 1993 GACACGAGCTGATCAACCGGCTTAAAGCGGAAATTTCAATTTCAAGGCGGCGAG 2052
 QY 2049 GCGGAGTGTGCGCAATTTCTTCAATGAGGAAATTTGACAGTACAGATTAATGCA 2108
 DB 2053 GCGGAGTGTGCGCAATTTCTTCAATGAGGAAATTTGACAGTACAGATTAATGCA 2112
 QY 2109 AATGCAATTTGAGTGTGTAAGATGTAATTCACAGAGCGCAATTTGTGTG 2168
 DB 2113 CAAGAGATTTTGTGTGTGCGACCGCATCAAGGCAATTCAGATGTAACGTTGAGCTGG 2172
 QY 2169 ACGATTTAACGATTTGTCAAAAAGTGAATTTAACGATTAACGATTAATTTCTA 2228
 DB 2173 ACGGCTGAGCAATTTGTGTGCAAAAACATTAACGAGATTAAGTATGCTCATTTG 2232
 QY 2229 CCAAAAACCAAAATCAATGCTCTATTAATTTAATGATTAACAGGCGCAATGTTAA 2288
 DB 2233 ACTTAACCGCATGACGAGCAATGCTGAGCTTGCAGTACAGCTCATTTAATCTCA 2292
 QY 2289 GGTTAAGCAAACTTATGCAATGCTATTAACAAATCAAGCAATTTAATCAATTAAC 2348
 DB 2293 GGGCTTGCACATCAACGCAATCTTATGTAAGGAGTACAGCTTAAACGTCAGC 2352
 QY 2349 AACATGCAACCAATTAAGCAATTTG----- 2377
 DB 2353 CACAAAGCAACCAAAACGCAACCTTATGCTGTGGCAATGCCCAAGCAATTTAAT 2412
 QY 2378 ----- 2377
 DB 2413 CAAGCAATTAACGCAACACATGCGCTTGGGCAATGCTTCAATTTAATCTTAAAC 2472
 QY 2378 -----ACTTCCGCAATTTGAATGCAACGTTGATTAAT 2411
 DB 2473 AACCGCGCAAAAGGCGAGTGTGAGCTTTCCGACAGGCTTAAGCAACGTAACCAT 2532
 QY 2412 GCAACTTGAAGGCTATGTCATTTAAGATTAAGCTGCAATTTCTTAAAGAAC 2471
 DB 2533 TCGCACTCAACGCGCAATGCTTCCCTACCGATTAAGCGAGTATTCATTTTGAACAC 2592
 QY 2472 CATTTTGCACCAATTTGAGGAGACAAAGCAACAGTACGTTGAAATGCGACT 2531
 DB 2593 CGCTTTACCGACACACTGACGCGGAGAGATGAGCTTACATTTAAAGACGGA 2652
 QY 2532 TGGCAATGCTTACGATTAATTTGCAATTTAAACGATTAATTAACGATGAGTACG 2591
 DB 2653 TGGACGCTGCGGTACGAGCGAATTAAGGCAATTTAACTTGAACAGCGCAATTTAC 2712
 QY 2592 TTAATTTGAGCTTAT-----TCAGTACGCTAAACATTAAG 2627

Db 2713 CTCGATTCGGCTATTCGACGATGCTGAGGCGCGCAACCGGCGATGCTGACAGCG 2772
 Qy 2628 CCAAGTCGCGCGTTCAT-----TAGAGACGAGAAACAAGCCAAACATCGGAGAAACAT 2678
 Db 2773 CCGCGCGCGCGTTCGCGCGCTTCCTATTATTCGTAACCGCGCAACTTCGTAAGATCC 2832
 Qy 2679 CGTTTCACACATGAGAGTAATAGTAATAGTGGGCAAGGCAATTCGTAATAC 2738
 Db 2833 CGTTTCACACGCTGAGGTAAGGCAATAGGCAATGAGGCAATTCGCTTTAG 2892
 Qy 2739 TCATCTTATTTGGCTATTAAGCGTAATTAATTAATTCATGTAAGCTGAGGCGAT 2798
 Db 2893 TCGGAATCTTCGCGCTACCGAGCGACCAATTTGAAGCTGCGGAAGTTCGAGGCGACT 2952
 Qy 2799 TACATATATCTGTTCCGACACAGGCAAGACCGCAACCGCTTGAGCATTAACCTTG 2858
 Db 2953 TACACCTTGGCGGTAAACATACCGCAACGAAACCGTAAGCTGATTCATTAACGCTA 3012
 Qy 2859 GTTGAAGCAAGATATCAACCGTATCAAGTAAGTCAATTTACTTTAGAAATGAC 2918
 Db 3013 GTGGAAGGAGAAAGCAACAAACCGCTGCGAAACCTTAATTTACCGCTGAAACGAA 3072
 Qy 2919 CAGTTGATGAGGTGATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAAT 2978
 Db 3073 CAGTGTATGCGCGCGGTGCGGTTCACACTACCGCAAGAGCGGAGTTCGCGCTG 3132
 Qy 2979 CATACCCCAATTAAGAGAGAAATGCAACATGATTTAGTAAGAGAGAGAGAGAA 3038
 Db 3133 CATATCCGCTGTAAGAAAGCAAGCTTTCCGCAACACTCGGCAAGGCAAGAACCGA 3192
 Qy 3039 CGAATTTAGAGAGCA---ACAAGTGAACCGACTCTTAACCAACAAACAGTGAACCA 3095
 Db 3193 CAGCGGAGAAAGCAACGCGCAACCGCTGACGCGTATTCGCGCGCGGAGGAGCC 3252
 Qy 3096 AAGTCGCGTCAAG 3155
 Db 3253 GCCGAAAG 3312
 Qy 3156 CTGTAAAGCATTAAG 3215
 Db 3313 ATTATCGAGCGAG 3354
 Qy 3216 GCAAAAAAGTTCGAGTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3275
 Db 3355 -----AAG 3406
 Qy 3276 AGCTGTGCTATTAAG 3335
 Db 3407 CCGCTTCCCG 3459
 Qy 3336 GATGCTCTAGCTCAAG 3395
 Db 3463 CAGCGCAACCTCAACCCCAACG-----CAGCGCACTGATGAGCGCT 3504
 Qy 3396 TATTCAAATAGTGCCTTATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3455
 Db 3505 TATGCAATAGGAGGTTGAGTAATTTCCGCAAGCTCAAGAGAGAGAGAGAGAGAG 3564
 Qy 3456 GATGAATAGATGCTCTTTTGTAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3515
 Db 3565 GACGAATGAG 3624
 Qy 3516 CAGGATTAAG 3575
 Db 3625 CCGGAGACCAAG 3681
 Qy 3576 TTAAGTCAATTTGAG 3635
 Db 3682 CTGCGCAATGAG 3741
 Qy 3636 CATAGCGCTTCAAGATTAATCTTTGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3695
 Db 3742 CACAAACGAG 3801

Qy 3696 ATGTGCGGTTTCCCGCAATATCAATGAGGAGATTAATTAATTTGATGAACGTGAGAACG 3755
 Db 3802 GCGCGCGGTTTCGAGCAATAGGAGATGCGAGAGGTTGACATCGGCATACAGCGCGCG 3861
 Qy 3756 GGAATGAGTGCAGTAATTAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3815
 Db 3862 GGTATTAG 3921
 Qy 3816 AATTATGCGAG 3875
 Db 3922 CATAGGAGATTAAG 3981
 Qy 3876 TTGAGATTAATGAG 3935
 Db 3982 ATCGGCGAG 4041
 Qy 3936 AAGCGCTGAGCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3995
 Db 4042 GCGACCGCGGCTTCGTTCAACGCGTACCGCGCGGAGATTAAGGAGAGATTAATTAAT 4101
 Qy 3996 ACTCGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4055
 Db 4102 AAGCGAG 4161
 Qy 4056 AAGCGTACGTAACAAACAGAGTAATTCACGAGTTCGAGAGAGAGAGAGAGAGAGAG 4115
 Db 4162 TCGGCGAG 4221
 Qy 4116 TCGCAAAAG 4175
 Db 4222 CCGATGCGAGATGAGGCGTAAACCGGAAATCAAGGTTTCAGCGTCTCCGCAAGCT 4281
 Qy 4176 TCAAAATCTAAGGTTCAACATCGGCAACAGCAAAATGAGGAGAGAGAGAGAGAGAG 4235
 Db 4282 GCGCGCGAG 4341
 Qy 4236 CGTTGTTAA 4244
 Db 4342 CCGCTGTAA 4350
 RESULT 10
 AA212252
 ID AA212252 standard; DNA; 4350 BP.
 XX
 AC AA212252;
 XX
 DT 08-OCT-1999 (first entry)
 XX
 DE Neisseria meningitidis strain A complete ORF1 sequence.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 XX treatment; Neisseria infection; meningitis; septicemia; gonorrhea; ss.
 OS Neisseria meningitidis.
 PN W09924578-A2.
 PD 20-MAY-1999.
 PF 09-OCT-1998; 98MO-IB001665.
 PR 06-NOV-1997; 97GB-03023516.
 PR 14-NOV-1997; 97GB-00024190.
 PR 18-NOV-1997; 97GB-00024386.
 PR 27-NOV-1997; 97GB-00025158.
 PR 10-DEC-1997; 97GB-00026147.
 PR 14-JAN-1998; 98GB-00000759.
 PR 01-SEP-1998; 98GB-00019016.
 PA (CHIR-) CHIRON SPA.
 XX

PI Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;
 XX WPI: 1999-327407/27.
 DR P-PSDB; AAY38824.
 XX
 PT Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for
 PT diagnosis, treatment and prevention of infection.
 XX
 PS Claim 9; Page 364-365; 524pp; English.
 CC
 CC Nucleotide sequences A4211972-212358 represent open reading frames (ORFs)
 CC of *Neisseria meningitidis* and *N. gonorrhoeae* which encode antigenic
 CC proteins (see AAY38499-Y38944). The antigenic proteins, their fragments,
 CC their nucleic acids and antibodies are used for diagnosis, prevention (as
 CC vaccines) or treatment of *Neisseria* infections, such as meningitis,
 CC septicemia and gonorrhea. Both organisms are closely related. Fragments
 CC of the nucleic acids are useful as hybridisation probes and antisense
 CC reagents
 XX
 SQ Sequence 4350 BP; 1225 A; 1161 C; 1060 G; 842 T; 0 U; 62 Other;
 Query Match 31.2%; Score 1348; DB 2; Length 4350;
 Best Local Similarity 59.0%; Pred. No. 0;
 Matches 2551; Conservative 0; Mismatches 1565; Indels 210; Gaps 9;

81 CTTAATTTTAAACCGCTTCATTTCAATAGGATAGTATGCAAGCGTGGCTGGTCAAC 140
 73 CCGCTACTTACGCAATATGCTGTGGTTCGCAATTTTCCCAACCTTGGCGGACAC 132
 141 ACTAATTTGGGATGATTACCAATATATCTGTATTTGGCGGAATAAAGGAATTC 200
 133 ACTAATTTGGGATGATTACCAATATATCTGTATTTGGCGGAATAAAGGAATTC 192
 201 AAGTGGGGCTCAAAATATTAAGTTTAAACAAAGGCGCAATAGTGGCAATCA 260
 193 GCAGTGGGGCTCAAAATATTAAGTTTAAACAAAGGCGCAATAGTGGCAATCA 252
 261 ATGCAAAAGCCCGCATGATGATTTTCTTATGTCAGTAAAGCGCTGGACGCTG 320
 253 ATGCAAAAGCCCGCATGATGATTTTCTTATGTCAGTAAAGCGCTGGACGCTG 312
 321 GTTAAATCATATATTTGTGAGGTGGCAATTAAGGATATCAATGTTGATTT 380
 313 GTTGGCATATATATTTGTGAGGTGGCAATTAAGGATATCAATGTTGATTT 372
 381 GTTGGCATATATATTTGTGAGGTGGCAATTAAGGATATCAATGTTGATTT 440
 373 GTTGGCATATATATTTGTGAGGTGGCAATTAAGGATATCAATGTTGATTT 432
 441 AACTCAAAAGGATATTTTATCATCTTATGAGACATTAACATATCAAGATTAAT 500
 433 AATTATAGCTGACATTTTCAACCTTCAACAGGATATNANATGCGGTTTCAT 492
 501 AATTGCTTACGAGACGGCTTCATATGATATGATCTTGAATATGATGATGAT 560
 493 AATTGCTTACGAGACGGCTTCATATGATATGATCTTGAATATGATGATGAT 552
 561 TCAGATGAGCAAAATATTCAGAAAGGTTCGATGAGCTTGGAGCGGATTTGGGA 620
 553 TCAGATGAGCAAAATATTCAGAAAGGTTCGATGAGCTTGGAGCGGATTTGGGA 612
 621 AATGATCAAGCAAAAGGATGAGCAAGGTTGCGGTGATATCATATCTGACAGCTGG 680
 613 TATGATGAGCAAAAGGATGAGCAAGGTTGCGGTGATATCATATCTGACAGCTGG 672
 681 AATGATCAAGCAAAAGGATGAGCAAGGTTGCGGTGATATCATATCTGACAGCTGG 740
 673 AATGATCAAGCAAAAGGATGAGCAAGGTTGCGGTGATATCATATCTGACAGCTGG 732
 741 GCGGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 800
 733 GCGGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 792

801 TTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 860
 793 TTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 852
 861 CTTTATGAGCAAAAGGATGAGCAAGGTTGCGGTGATATCATATCTGACAGCTGG 917
 853 CTTTATGAGCAAAAGGATGAGCAAGGTTGCGGTGATATCATATCTGACAGCTGG 912
 918 ATTGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 977
 913 ATTGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 972
 978 ATTGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1037
 973 TTTATGAGCAAAAGGATGAGCAAGGTTGCGGTGATATCATATCTGACAGCTGG 1032
 1038 ATTGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088
 1033 CCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1092
 1085 GTTCAATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1148
 1093 GAACCAATTTACGGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1152
 1149 AGGCTATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1208
 1153 AACTTTCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
 1209 GGGGGGGGGTCTTATTTTGAAGGATGATGATGATGATGATGATGATGATGATGATGAT 1268
 1213 GGGGGGGGGTCTTATTTTGAAGGATGATGATGATGATGATGATGATGATGATGATGAT 1272
 1269 TGGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1328
 1273 TGGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1332
 1329 GTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1388
 1333 GTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1392
 1389 GAAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1448
 1393 GAAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1452
 1449 GATCAAGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1508
 1453 GATCAAGCAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1512
 1509 CAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1568
 1513 CAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
 1569 CGCTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1628
 1573 CGCTTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1632
 1629 GCAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1688
 1633 GCAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1692
 1689 AGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1748
 1693 AGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1752
 1749 TACCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1808
 1753 TACCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1812
 1809 TATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1868
 1813 TATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1872
 1869 GATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1928

```

1873 AATGACGCAACAAAGCGAAGCTTTTTCAGCGGAGACGACACCGCAGCTC 1932
1929 AATCATTTAATTAACCTTGTGAGAAATGAGAGATACACAGCGGAAATTTGTG 1988
1933 AATCATTTTGAAGCGGTGTGTCAAAAATGCAAGGTATCCCAACAGGAAATTTGTGTG 1992
1989 GATCAGATTGATACACCTGATCTTAAGCTGAAACCTTCAATTTAAAGGCGAAGT 2048
1993 GACAGAGCTGATACACCGACGTTTAAAGGAAATTTCCATTTTCAAGGCGAG 2052
2049 GCGGTGTTTCTCGCAATGTTTCTCAATTTGAGGAAATTTGACAGTACGCAATATGCA 2108
2053 GCGGTATTTCCCGAATTTTTCGCAAGTGAAGGCGATTTGATTTGAGCAATACGCC 2112
2109 AATGCCAATTTGTTGTGTCGCAAAATCAAAATACATTTGACCGCTTCAGATTG 2168
2113 CAAGCAGTTTGTGTGTGCGACCGCATCAAGCCATACATCTGTACACGTTCCGACTGG 2172
2169 AAGGATTACGACTGTGCAAAAATGCAATTTAACGATACAAAATTTAATCTATA 2228
2173 ACGGTTGACAAATTTGTGTGAAANAACATTAACGATTAAGTATTTCTCATTTG 2232
2229 CCAAAAACACAAATCAATGCTCTATTATTTAACTGATTAAGCAACGCGCAATGTTAA 2288
2233 ACTAAGACGACATTAAGGCGCANTGTACGCTTTCACGATTAAGTATTTCTCATTTG 2232
2289 GGTTTAGCAAACTTAATGCAATGCACTTTAAACAATACAGCCATTTACTTAAGC 2348
2293 GGGCCTGCACTTAATGCAATCTTAAGTAAATGCGATACAGTTATACGTCAGC 2352
2349 AAGATGCCACCCAAATGAGCAATTTG 2377
2353 CACAGCGCACCCAAACGCGACCTTAAGCTGTGTGCAATGCGCAACGACCTTAAT 2412
2378 ----- 2377
2413 CAAGCAGCTTAAGCGCAACGCAATGCGGCAATGCTTCAATTAATCTAAGCAAC 2472
2378 ----- 2411
2473 AAGCGCGCACAAACGCGAGCTGACGCTTTTCCACAAAGCTTAAGGCAAGTAAAGCCT 2532
2412 GCAAACTGGAAGGATATGATTAAGATTAAGATTAAGTCAATTTCTTTAAAGAAAGC 2471
2533 TCCGCACTCAAGGCAATGCTCCCTGACGATTAAGGCAATTTCAATTTGAAAGAGC 2592
2472 CATTTTTCGCACAAAATTTCAAGGAGCAAAAGCGCAACAGTGAAGTTGAAATGCGACT 2531
2593 GCGTTAACGCGCAACTAGGCGGACAGAGAAAGCATTAAGTAAAGACAGGAA 2652
2532 TCGACAAATGCTAGAGATCTAATGTCAGATTTAAAGCTTAATTAAGTACATGAG 2591
2653 TGAACGCTGCTGAGGACGGAATTAAGCAATTTAAACCTTGAACAGCGCAACATTAACA 2712
2592 TTAATTTAGCTTAT-----TGAAGTACGCTCAAAAGCAATAGC 2627
2713 CTCAATTTCCGCTTATGCGACGATGCTGACGCGGCAACCGGCAAGGTCTGAGACG 2772
2628 CCAAGTCCGCTTCA-----TGAAGCGGAAACAAAGCGCAACATGCGCAAGAAAT 2678
2773 CCGGCGCGCGCTTCCGCGCTTCCATTAATCCGTTACACGCGCAACTTCCGTTAGATCC 2832
2679 CGTTTCAACATTAAGCAATTAATGATTTGAGTGGGAAAGGCAATTTCACTTTACT 2738
2833 CGTTTCAACAGCTGACGATTAAGCGCAATTTGAACNGTCAAGAAACATTTCCGTTTAG 2892
2739 TCACTTATTTGCTAATTAAGCAATTAATTAATTTCACTGAGCTGAGGCGCAT 2798
2893 TCGGAACCTTTGCGTACGAGACGCAATTTGAAGCTGCGGAAAGTTTCCGAAAGAACT 2952
2799 TACATTTATCTGTTGCAACAGGCGAAAGAACCGGAAACCTTGAAGCAATTAATTTG 2858

```

```

2953 TACACTTGGCGGTCAACATTAACGCGCAACCGGTAAGCTGATCACTGAAGGTA 3012
2859 GTTGAAGCAAGATTAATCAACGTTTACAGTAAGCTCAATTTACTTTAGAAATGAC 2918
3013 GTGGAAGGAAAGACAAACAAACCGGTGTCCGAAATCTTAATTTACCTTCAAAACGAA 3072
2919 CACGTTATGACAGGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2978
3073 CACGTGATGACGCGCGCGGTGTGTTACCACTATCCGAAAGAGCGGATTCGCGCTG 3132
2979 CATTAACCAATTAAGAGCAAGAAATTCACATGATTTAGTAAGCAAGCAAGCAAGA 3038
3133 CATTAATCCGCTCAAAAGAACAGAGCTTTCCGACAACTGCGAAGGC----- 3179
3039 CCAACATTTGAAGCCAAACAGTGAACGATCTTAACCAACAAACAGGTGACCCAAA 3098
3180 -----AGAACCAAAACAGGCGGAAAGAACAGGCGCAAAACCTTTGAGGCTG 3221
3099 GTGCGGTCAAGAGACAGAGACAGCGCTTCTGATACCTGCTGATCAAAAGCTG 3158
3232 ATTTGCGCGCGCGCGCATGTC---CGCGAAAGACAGAAAGCTTTGCGAACCAGCGCG 3288
3159 TTAACGCTTAAGAACCCAAACAGCTGAACGCTGCTGAACACAAAAAGTAAGCA 3218
3289 CAGCAGCGCGGCAAAATGTCGACATTAACAGCGCGAGAGAAACAGCGGTGACG 3348
3219 AAAACAAAAAGTGGCGGTCAAAAGAGCAGTGTTCGATCCCTGCTGATCAAAAGC 3278
3349 GCGGATTAAGCAGCGCCTTTGCGAAACAGCGAAAGCGGAAACCGCGGATACAC 3408
3279 CTGTTGCACTTAAGACCGCACTTGAAGTTATGATGCCCCACAGCAATGCAAAAGAT 3338
3409 GCGTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3458
3339 CGCTAGCTTAAGAGAGAGCGGAAACAAACGCAACAAACAAACAAACAAACAAACAAAC 3398
3469 CCTCAACCCCA-----CGCAGCGCGACCTGATTAAGCTTAT 3507
3399 TCAATTAAGCTTATCAAGATTAATCTGCAACAGTAATTAAGTATGCTTGTCAAGT 3458
3508 GCGAATAGCGGTTGAGTAAATTTCCGCAAGCTCAACAGGTTTCCGCAAGAGAC 3567
3459 GAATTAAGCTGCTTTTGTAGATCAAGCAACATCTGCGGTGTGAGCAAAATGCAAGC 3518
3568 GAATTAAGCGCGGTTTCCGCAAGCGCGCAAGCGGTTTGAACAGCAGCATCCG 3627
3519 GATTAAGAGCGCTATGATTTGATGCTGCTTCCGCTTATCAAGCAAGCAAAAGAACTTA 3578
3628 NAACCAAAACATACGCTTCCGCAAGATTTCCGCGCTTACCGCAAC---AAACCGACCTG 3684
3579 CGTCAATTTGGGCGGCAAAAGCCTTATGATGACGAATTTGGGCGAGTTTCTGCAAT 3638
3685 CCGCAATGCTATGAGAAACCTTGCAGCGGCGGCGGTGCGATCTGTTTTCGAC 3744
3745 AACCGGACGAAACANCTTTCAGACGCGCATCGCACTGCGCAACGCTTCCGCAAGC 3804
3699 TCGGCTTTGCGCAATTAATGATGAGGCGATTTCAATTTGATGTAACGTTGGAAGGGA 3758
3805 GCGGTTTTCGCGCAATTAAGGCAATGCGGATTTGAGATGCGATGACAGCGCGGCT 3864
3759 ATCAGTGCAGTAAATGCTGAGAGACAAAGCCGAAATTCATGCAAAAGGATTAAT 3818
3865 TTAAAGCGGCAATCTTATCAAGCGATGAGAGGAAATCCCGCGCGCGCTGCAAT 3924
3819 TATGCGGTGAATCAAGTATCAAGTTCCGTTTGAAGCAATTTGAGCAATCAAGCTTATTT 3878
3925 TACGCAATTAAGGCAAGATACGCGCGGTTTGGGAGTTTCCGCAATGCAAGCTGATC 3984
3879 GAGATTAATGCTATTTTATGAAGTGAATAATCAACTGAGGAAGGAGAGGAA 3938
3985 GCGGCAAGCGCTATTTGCTCAAAAGCGAATTAACCGCTACGAAACGTCATATGCGC 4044

```


QY 3939 ACGGTAGCCTTSCATTTATGCTATATGCTGCGATTCGATGATTAATACCTTACT 3998
 DB 4045 ACCCCCGCTTTCGTTACCGTACCGGCGGATTAAGGAGATTTATTCATTTCAA 4104
 QY 3999 CCGACAGATTAATTCAGCGTTAGCCTTATTTCTTCTGCAATTAATGATGTTTCAAC 4058
 DB 4105 CGGCGGACAGATTTCCATTCACGCTTATTTAGCGTCTGCTTACCGATCGGCTTCG 4164
 QY 4059 GCTAACCTACAAACACGCTTAATTCACGCTGTTGCAACACCTTTGACGTTATGG 4118
 DB 4165 GCGAAATCCGAAACAGCGCTCAATACCGGATTTGCTCGAATTTCCGCAAAACCCG 4224
 QY 4119 CAAAAGAGTGGGATTAAGGCAAGAAATTTTACATTTCCAAATTCGCTTTATCTCA 4178
 DB 4225 AGTGGGATGCGGCGTAAACGCGGATTCAGATTCAGGTTTACGCTGCTTCACCGCTCC 4284
 QY 4179 AATCTCAAGCTTCACACTGCGCAACAGCAAAATGCGGCGCTGAATTTGCGTATGCT 4238
 DB 4285 GCGCGCAAAAGNCCGCACTGGAAGCCCAACACAGCGCGGCAATTAATTAAGCTACCG 4344
 QY 4239 TGGTAA 4244
 DB 4345 TGGTAA 4350

RESULT 11

AAZ12253

ID AAZ12253 standard; DNA; 4407 BP.

AAZ12253;

08-OCT-1999 (first entry)

Neisseria gonorrhoeae complete ORF1 sequence.

Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

treatment; Neisseria infection; meningitis; septicemia; gonorrhea; ss.

Neisseria gonorrhoeae.

MO3924578-A2.

20-MAY-1999.

09-OCT-1998; 98MO-IB001665.

06-NOV-1997; 97GB-00023516.

14-NOV-1997; 97GB-00024190.

18-NOV-1997; 97GB-00024386.

27-NOV-1997; 97GB-00025158.

10-DEC-1997; 97GB-00026147.

14-JAN-1998; 98GB-00000759.

01-SEP-1998; 98GB-00019016.

(CHIR-) CHIRON SPA.

Masigiani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;

WPI; 1999-327407/27.

P-PSDB; AAY38825.

Proteins from Neisseria meningitidis and N. gonorrhoeae useful for

diagnosis, treatment and prevention of infection.

Claim 9; Page 370-371; 524pp; English.

CC of the nucleic acids are useful as hybridisation probes and antisense
 CC reagents
 XX
 SQ Sequence 4407 BP; 1273 A; 1169 C; 1094 G; 871 T; 0 U; 0 Other;

Query Match 30.0%; Score 1297.8; DB 2; Length 4407;
 Best Local Similarity 59.1%; Pred. No. 6,4e-303;
 Matches 2571; Conservative 0; Mismatches 1567; Indels 213; Gaps 12;

QY 86 TTTTAAACCGCTTGCAATTCATTAAGGATTAATGCAAGCTGGGTGCTCACTTA 145
 DB 78 TTACTTACGATTAATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 137
 QY 146 TTTTGGATTAATTAACCAATTAATGATTAATGATTAATGATTAATGATTAATGATTA 205
 DB 138 TTTTGGATTAATTAACCAATTAATGATTAATGATTAATGATTAATGATTAATGATTA 197
 QY 206 TGGGCTCAAAATTAATTAAGTTTAACCAACAAAGGCAATTAATGATTAATGATTA 265
 DB 198 CGGCGCAAAAGTATTAAGTTTAACCAACAAAGGCAATTAATGATTAATGATTAATGATTA 257
 QY 266 AAAAGCCCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 325
 DB 258 GAAAGCCCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 317
 QY 326 AAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 385
 DB 318 CGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 377
 QY 386 AGAGGAAACAAACCCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 445
 DB 378 GAGGAAACAAACCCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 437
 QY 446 CAAAAG---KAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 502
 DB 438 TAAAGCAAGGACTTAACCGCATCTTAATGCGGATTAATTAATTAATTAATTAATTAATTA 497
 QY 503 ATTGCTTACAGAGCGGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 562
 DB 498 ATTGCTTACAGAGCGGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 557
 QY 563 AGATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 622
 DB 558 TGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 617
 QY 623 TGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 664
 DB 618 TGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 677
 QY 665 TCGACAGCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 724
 DB 678 GCTGTGCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 737
 QY 725 AGGCAATGTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 784
 DB 738 TAGGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 797
 QY 785 CAGTGTGCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 844
 DB 798 CAGTGTGCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 857
 QY 845 ATTAAGGAAACAAACCCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 904
 DB 858 ATTAAGGAAACAAACCCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 917
 QY 905 TTAATTT---TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 961
 DB 918 TTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 977
 QY 962 TTAATGAGTGAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1013
 DB 978 AATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1037

QY 1014 CAGAAATCAGAAATCCATGAGAAATTAATTAAG-----TTAGCAATATGAG 1063
 Db 1038 TAAAGCTATCTCTACCTTATGATTAATAAAGCGCTCAATGTTTAAATGTTTC 1097
 QY 1064 TTTACCTTGAAGAGAGATTAAGTTCAATACCTTAAGATGACGACCTAATATTTA 1123
 Db 1098 TTTATCCGAGACGAGAGAACTGTTTATATGCTGCA--GGTGGGCTCAAGATT 1154
 QY 1124 TTTCAAGCTTTAAACATGAGAGAAAGCTATATTTTATGATCAAAAACAAGATCAT 1183
 Db 1155 TCGACCAAGCTGAATTAATGAGAAATATTTCTTTATGACAAAGAGAAAGGATTT 1214
 QY 1184 AATCTCGATCTGACATTAACAAGGCGGCTGCTTATTTTGAAGGTAATTTTAC 1243
 Db 1215 GATCTTACCAACATCAACCAAGGCGGCGCTTGTATTTGAGGTAATTTTAC 1274
 QY 1244 AGTATCTCCAAATTTTAACTTGGCAAGGCTGAGTACATGTAAGTAATG 1303
 Db 1275 GGTCTGCTTAAACAAAGAAAGTGGCAAGGCGGCTCATATCATGATGAGCAG 1334
 QY 1304 CACGTTACTTGAAGTAATGCGCTGGAACATGATCGACTTTCTAAATTTGTAAAG 1363
 Db 1335 TACGTTACTTGAAGTAATGCGCTGGAACAAAGCGCTGCTCAATTTGCAAAAG 1394
 QY 1364 AATCTGACGTTCAAGCCAAAGGAGAAATTAAGCTTCAATGAGCTGATGCTAA 1423
 Db 1395 CAGCTGCTGCTTCAAGCCAAAGGAGAAATTAAGCTTCAATGAGCTGATGCTAA 1454
 QY 1424 AGTATTTTGGACGAGCGAGACATCAAGGCAAAAGCTTATGTAAGTAATG 1483
 Db 1455 AGTATTTTGGACGAGCGAGACATCAAGGCAAAAGCTTATGTAAGTAATG 1514
 QY 1484 CTGCTGAGCGGAGAGGAGCTGCTTCAATTAACGATTAACAATTTGATCCGATA 1543
 Db 1515 CTGCTGAGCGGAGAGGAGCTGCTTCAATTAACGATTAACAATTTGATCCGATA 1574
 QY 1544 ATTTATTTGGCTTGGTGGTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 1603
 Db 1575 ACTCTATTTGGCTTGGTGGTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 1634
 QY 1604 ACCTATTTGGCTTGGTGGTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 1663
 Db 1635 CCGATTTGGCTTGGTGGTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 1694
 QY 1664 TATGCTATTTGGCTTGGTGGTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 1723
 Db 1695 CACGCTTACCTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 1748
 QY 1724 ACTGATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 1783
 Db 1749 CTGATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 1808
 QY 1784 ACATATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 1843
 Db 1809 AACATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 1868
 QY 1844 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 1903
 Db 1869 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 1928
 QY 1904 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 1963
 Db 1929 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 1988
 QY 1964 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2023
 Db 1989 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2048
 QY 2024 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2083
 Db 2049 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2108
 QY 2084 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2143

Db 2109 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2168
 QY 2144 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2203
 Db 2169 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2228
 QY 2204 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2263
 Db 2229 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2288
 QY 2264 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2323
 Db 2289 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2348
 QY 2324 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2383
 Db 2349 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2408
 QY 2383 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2468
 Db 2409 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2483
 QY 2447 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2506
 Db 2469 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2528
 QY 2483 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2546
 Db 2509 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2568
 QY 2547 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2583
 Db 2569 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2606
 QY 2604 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2668
 Db 2629 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2683
 QY 2664 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2723
 Db 2689 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2748
 QY 2724 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2783
 Db 2749 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2808
 QY 2784 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2843
 Db 2809 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2868
 QY 2844 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2903
 Db 2869 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2928
 QY 2904 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2963
 Db 2929 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 2988
 QY 2964 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 3023
 Db 2989 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 3048
 QY 3024 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 3083
 Db 3049 CAAATTTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 3108
 QY 3084 TTTAGTGTGTAATTTAAAGGCTTGGCTTAACTTAAAGGCTTCAATTTGATCCGATA 3143

Db 3189 ACTGCGAAGCGGAGGAAAGAGAGCCGCTTGAACGACAAACAACTTCGCGC 3248
 QY 3074 TAAATACAAACAGGTGAGCCAAAGTCCGTCAAGAGAGAGAGAGAGGCTTTC 3133
 Db 3249 CAAACAAACGCGGAAAAAGACAGCGCAAGCCTTGACGCTGATTTGGCGCGG 3308
 QY 3134 TGATACCTGCTGATCAAGCCCTGTAAAGCATAGAAAGCCAAACAGCTGAATGAC 3193
 Db 3309 CATGTGCAACGAAAGGCAAGAGGTGTCCGAAACCGCGCGGCGAGGCGGAAAA 3368
 QY 3194 TGTGAACACAAAGAAAGTAAAGCAAAACAAAGAGTCCGCTCAAAAAGACAGTGT 3253
 Db 3369 TGCCGCAATTATGACGCGGAGAGAGAGAAAAAGCGGTGACAGCGATTAAGACACCG 3428
 QY 3254 TTCTGATCCCGCTGATCAAGAGCTGTGACATTAAGAGCGGACCTTGAAGTATTA 3313
 Db 3429 CTGCGGAAACGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3488
 QY 3314 TGCCCAACAGCAATGAGAAAGAGTGTACCTCAAGAGAGAGAGAGAGAGAGAG 3373
 Db 3489 CCGGCGCGCGCGGATTTGCGGCAACCGAGCCCAACCGCAACCGCAACCGGCGC 3546
 QY 3374 ACAGAAAGCTTGAATGACCGGTATTAACAAATAGTCCGTATCAAGATTTGCAACAGT 3433
 Db 3547 -----GACCTGATCAGCGGTATGCAATAGCGGTGATGATTTTCGCGACGCT 3599
 QY 3434 AAAGATGATCTTCTGTCAGAGTATTAAGTGTCTTCTTGTATGATCAAGCAATC 3493
 Db 3600 CACAGGCTTTTCCGCTTACAGAGATTAAGTGTCTTCTTGTATGATCAAGCAATC 3659
 QY 3494 TGCGGTGAGCAAAATATGCAAGAGATTAAGAGAGTATGATTTGATTTGCTGCTG 3553
 Db 3660 CGCGTTTGAACAGCGGCAATCGGAGACAGAAACAGTACGTTGCAAGATTTCCGCGC 3719
 QY 3554 TTATCAGCAGCAAAAGAACTTACCTCAATTTGGGCTGCAAAAGCTTACTATG 3613
 Db 3720 CTACCGCAAC---AAACGACCTGCGCAATGCTATGCAAAAAACCTGCGACGCG 3776
 QY 3614 ACGAATTTGGGAGGTTTCTGCAAGTACGCGCTTCAATTAATCTTGTATGAACAGTTA 3673
 Db 3777 GCGGTGCGATCTGTTTGTGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3836
 QY 3674 AATACGCGAGATTAACAGATGATGATGATTTTCCCAATATCAATGAGGAGATTA 3733
 Db 3837 CAATGCGGACGCGCTTGCACAGGTGCGCTTTGCGGAAATAGGCAATGCGAGTTTGA 3896
 QY 3734 ATTTGCTTAAAGTGGAGAGAGAGATGATGCGATTAATGCTGAGAGAAACAGCG 3793
 Db 3897 CATGCGATCAGCGCGGCGGCTTTAGTACGCGAGCCCTTCAAGAGGAGATGAGG 3956
 QY 3794 AAAAAATCATCGAAAGAGATTAATTAAGGCTGAAATGATTAATCAATTCGTTAG 3853
 Db 3957 CAAATCCGCGCGCGCGCTGCTGATACGCAATGAGCAATGAGGAGAGGTTTCG 4016
 QY 3854 GCAATTTGGGATTCAGCTTATTTTGAAGTATGCTATTTATTAAGTGAATA 3913
 Db 4017 CGAATTTGGGATTCAGCTTATTTTGAAGTATGCTATTTATTAAGTGAATA 4076
 QY 3914 TCAATCTGAG 3973
 Db 4077 CCGATGAGAAAGTCAATGCGACCCCGGCTTCAATCAACCGCTGACCGCGGAG 4136
 QY 3974 CATTCAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4033
 Db 4137 CATTAAGCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4196
 QY 4034 GGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4093
 Db 4197 CCGTCTTATACCGATGCGCTTCCGCAAGTCCGAAAGCGCTCAATGCGCTAT 4256
 QY 4094 GCAACACATTTGACAGTATTAAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 4153
 Db 4257 GCGGAGAGATTTGCGAAACCCGAGTGGGATGCGGATGCGGATTAAGTCAAGG 4316

QY 4154 TTTCGAATTTCCGCTTATATCAAAATCTCAAGTTTCAACTTCGCGAAACAGCAAA 4213
 Db 4317 TTTCAGCTGTCTTCTTCAAGCTGCGCGGCGGCAAGAGAGAGAGAGAGAGAGAG 4376
 QY 4214 TGTGGGCTGAATTTGAGCTATCGTTGTA 4244
 Db 4377 CCGGCGCATCAAAATTAAGCTACCGCTGTA 4407
 RESULT 12
 AB238249 standard; DNA; 4404 BP.
 ID AB238249 standard; DNA; 4404 BP.
 AC AB238249;
 XX
 DT 07-MAR-2003 (first entry)
 XX
 DE N. gonorrhoeae nucleotide sequence SEQ ID 1087.
 XX
 DE Antibacterial; infection; vaccine; gene therapy; gene; ds.
 XX
 OS Neisseria gonorrhoeae.
 XX
 PN W0200279243-A2.
 XX
 PD 10-OCT-2002.
 XX
 PR 12-FEB-2002; 2002MO-IB02069.
 XX
 PR 12-FEB-2001; 2001GB-0003424.
 XX
 PA (CHIR-) CHIRON SPA.
 XX
 PI Fontana MR, Pizsa M, Masignani V, Moraci E;
 XX
 DR WPI; 2003-058415/05.
 DR P-PSDB; ABP77279.
 XX
 PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 XX
 PT medicament for treating or preventing N. gonorrhoeae infection.
 XX
 PS Disclosure; Page 266-267; 815pp; English.
 XX
 CC The present invention relates to proteins from Neisseria gonorrhoeae.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records AB237705-AB242316 represent nucleic acid
 CC molecules of the invention
 XX
 SQ Sequence 4404 BP; 1271 A; 1170 C; 1094 G; 869 T; 0 U; 0 Other;
 Query Match 30.0%; Score 1294.8; DB 7; Length 4404;
 Best Local Similarity 59.1%; Pred. No. 3,4e-302;
 Matches 2568; Conservative 0; Mismatches 1567; Indels 213; Gaps 12;
 QY 86 TTTTAAAGCTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 145
 Db 78 TTACTTAAAGCTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 137
 QY 146 TTTTAAAGCTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 205
 Db 138 TTTTAAAGCTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 197
 QY 206 TGGGCTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 265
 Db 198 CCGGCGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 237
 QY 266 AAAGCGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 325

Db 258 GAAAGCCCCGATGATGATTTTCTGTGTATCGCCGTAACGGCGGCGCATTCGCGG 317
 Qy 326 AATTAATATATTTGTAGCGTGGCACAATAGTATGATATACAGATGTTATTTTGTGTC 385
 Db 318 CGATCAATATATTTGTAGCGTGGCACAATAGTATGATATACAGATGTTATTTGTGTC 377
 Qy 386 AGAGGGAAACAAACCCGATCAATATGTTTACTATTAAGATTTGTAACAAATATACTA 445
 Db 378 GAGGGAGACATCCGATGACACCGCTTTCTTACCAATTTGGAAAAATTAATTA 437
 Qy 446 CAAAAAG---ATAATTTATCTTTATGAGAGATTAACCAATATCCAGATTAACATA 502
 Db 438 TAAAGCAGGACTAACCGGCATCTTATAGCGGGAATTAATCATATGCGCGTTGACAA 497
 Qy 503 ATTCTTACAGAGCGCTCAATTAATGATCTTGAATATGATATGACATTAATTC 562
 Db 498 ATTGTCAAGATGAGAACCTGTGATGACCAATTAATGATATGATATGATATGAC 557
 Qy 563 AGATTAACAAATATGACAGATGATGATGATGATGATGATGATGATGATGATGAT 622
 Db 558 TGATTTAATTAATATCCCTGATGATGATGATGATGATGATGATGATGATGATGAT 617
 Qy 623 TGATCAGACAAAGGC-----GACCAATTTGCGGTGATATCATTA 664
 Db 618 TGATGAGACGAAACCAATACCGCAATTTCAATCATATATGCAAGCGATATCTTG 677
 Qy 665 TCTGACAGCTGGCAATACACATATAGGTGAGACGATATGATATGATATGAT 724
 Db 678 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737
 Qy 725 AGGCAATGTTGTAAGCGGAGATATGATGATGATGATGATGATGATGATGATGAT 784
 Db 738 TAGGAAAAAATTAATACATACCAATATGATTTTACCAAGAGGCTGATTTGCGCA 797
 Qy 785 CAGTGTCTCGATGATTTATATATATATATATATATATATATATATATATATAT 844
 Db 798 CAGTGTCTCGATGATTTATATATATATATATATATATATATATATATATATAT 857
 Qy 845 ATTAAGGAGGCAACCTTTTGAAGCAAAATATGATGATGATGATGATGATGATGAT 904
 Db 858 ATTGCAAGAGCAACCTTTATATGAAAAATGATGATGATGATGATGATGATGATGAT 917
 Qy 905 TTAATTT---TGATGAAATTTTCAAAAGATTTATATATATATATATATATATATAT 961
 Db 918 TTGATTTAT 977
 Qy 962 TAAATGATGATACAT 1013
 Db 978 AATGGAAT 1037
 Qy 1014 CAGAAATCAGAAAT 1063
 Db 1038 TAAACATATTTCTAT 1097
 Qy 1064 TTTACCTTTGAAAGAGATTAAGTATATATATATATATATATATATATATATATAT 1123
 Db 1098 TTTATCCGAGACGAGAGAACTGTTATATATATATATATATATATATATATATAT 1154
 Qy 1124 TTTCCACGTTTAAACATGAGAAACGATATATATATATATATATATATATATATAT 1183
 Db 1155 TCGACCCAGATGAT 1214
 Qy 1184 AATCTTGGATCTGATTTAAACAAAGGCGGTGCTTATATATATATATATATATATAT 1243
 Db 1215 GATCTTACAGCAAT 1274
 Qy 1244 AGTATCTCAAAATTTAAACAACTTGGCAAGAGCTGCTATATATATATATATATAT 1303
 Db 1275 GGTCTGCTTAAACAAACGAGAACTGCGAAAGCGCGCTTATATATATATATATAT 1334
 Qy 1304 CACCTTATCTTGAAGAAATGCGGTGAGAACTATGATCTTTTAAATTTGTAAGG 1363
 Db 1335 TACGTTACTTGAAGAAATGCGGTGAGAAACGCGCTTCTTCAAAATCGCGAAAG 1394

Qy 1364 AACATTCAGTTCAGCCAAAGGGGAAATTAAGTTTGATGACGGTAGCGATGTAA 1423
 Db 1395 CACCTGCTGTTCAAGGCAAAAGGGGAAATTAAGTTTGATGACGGTAGCGATGTAA 1454
 Qy 1424 AGTATTTTGAAGACGAGGAGATGATGATGATGATGATGATGATGATGATGAT 1483
 Db 1455 AGTATCTTATGATCAGCGAGGAGATGATGATGATGATGATGATGATGATGATGAT 1514
 Qy 1484 CTGTTAGGCGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1543
 Db 1515 CTGTTAGGCGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGAT 1574
 Qy 1544 ATTATTTTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1603
 Db 1575 ACTCATTTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1634
 Qy 1604 ACATATCCAAATATGAGAGGAGGAGGATGATGATGATGATGATGATGATGATGAT 1663
 Db 1635 CCGATTCGAAATATCCGATGAGGAGGAGGATGATGATGATGATGATGATGATGATGAT 1694
 Qy 1664 TAAATGCTATATCTGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGAT 1723
 Db 1695 CACCTTACATATACGAGGATTA-----GATATATCTAACCGGATATAGACAA 1748
 Qy 1724 ACTTATTTACGAAAGAAATTTGCTTCAACGGTTGCTTGGCAAAACAGATTAATTA 1783
 Db 1749 CTGATATGCAAAAGAAATTTGCTTCAACGGTTGCTTGGCAAAACAGATTAATTA 1808
 Qy 1784 ACATATGAGGAGGATTAACCTTATTTTAAACCAACGAGGATGATGATGATGAT 1843
 Db 1809 AACGAAACGGGAGGATTAACCTTATTTTAAACCAACGAGGATGATGATGATGAT 1868
 Qy 1844 TTCAGTGTATCAAAATTTAAAGGAGGATTAACCTTATTTTAAAGGATTAATTTT 1903
 Db 1869 TTCAGGAGGAGGATTAACCTTATTTTAAAGGAGGATTAACCTTATTTTAAAGG 1928
 Qy 1904 CCGTACAGGAGGAGGATTAACCTTATTTTAAAGGAGGATTAACCTTATTTTAAAG 1963
 Db 1929 CCGAGAGGAGGAGGATTAACCTTATTTTAAAGGAGGATTAACCTTATTTTAAAG 1988
 Qy 1964 TATACCAAGGAGGAGGATTAACCTTATTTTAAAGGAGGATTAACCTTATTTTAAAG 2023
 Db 1989 TATACCAAGGAGGAGGATTAACCTTATTTTAAAGGAGGATTAACCTTATTTTAAAG 2048
 Qy 2024 AATCTTCAAAATTTAAAGGAGGAGGATTAACCTTATTTTAAAGGAGGATTAACCT 2083
 Db 2049 AATCTTCAAAATTTAAAGGAGGAGGATTAACCTTATTTTAAAGGAGGATTAACCT 2108
 Qy 2084 AATTTGACAGTCAAGCAATATGCAATGCAATTTGCTTGTGCAATTTCAACAAA 2143
 Db 2109 CAGTTGAGGATTAAGCAATATGCAATGCAATTTGCTTGTGCAATTTCAACAAA 2168
 Qy 2144 TATACCTTGAAGGAGGATTAACCTTATTTTAAAGGAGGATTAACCTTATTTTAAAG 2203
 Db 2169 CAGATCTGTAAGGAGGAGGATTAACCTTATTTTAAAGGAGGATTAACCTTATTTTAAAG 2228
 Qy 2204 CAGATCTGTAAGGAGGAGGATTAACCTTATTTTAAAGGAGGATTAACCTTATTTTAAAG 2263
 Db 2229 CAGATCTGTAAGGAGGAGGATTAACCTTATTTTAAAGGAGGATTAACCTTATTTTAAAG 2288
 Qy 2264 TATATGCAAGGAGGAGGATTAACCTTATTTTAAAGGAGGATTAACCTTATTTTAAAG 2323
 Db 2289 CAGATCTGTAAGGAGGAGGATTAACCTTATTTTAAAGGAGGATTAACCTTATTTTAAAG 2348
 Qy 2324 AATTTCAAGCA----- 2355
 Db 2349 CAGAGCAGGAGGAGGATTAACCTTATTTTAAAGGAGGATTAACCTTATTTTAAAG 2408
 Qy 2336 ----- 2335
 Db 2409 GGGCAATGCCAAGCAATTTATGAGCAATTAAGGCAATTAAGGCAACATGCGCTTGG 2468

QY 2336 -----ATTACATTAAAGCAACATGCAACCAATAGCAATTCGACTTTCCGA 2386
 DB 2469 CAATGCTTCTTATCTTAAGCAACACCGCTGCAAAAGCGAGCTGACGCTTTCCGA 2528
 QY 2387 CAATTCACATGCAACGCTGATATGCAAACTTGAACGCTATATGCAATTTACGATTC 2446
 DB 2529 CAACGCTAAGCAAAAGCTAAGCAATTCGCACTCAACGCAATGCTCCCTAGCCATA 2588
 QY 2447 AGCTCAATTTCTTAAAAAGCCATTTTCCGACCAAAATTCAGGAGCAAAAGGCGC 2506
 DB 2589 GGCAGATTTCATTTTGAAGACGCGCTTACCGAAAAATACGGCGCAAGATAC 2648
 QY 2507 AACAGTAGCTTGAAGAAATGCACTTGAACATGCTAGGCAATCTACATTTGCAATTT 2566
 DB 2649 GGCATTACCTTAAAAAGCAAGCAATGACGCTGCTCGGAGCAAGATTTAGGCAATTT 2708
 QY 2567 AAGCTMAATACAGTACGATCAGTTAAATTCAGCTTAT----- 2606
 DB 2709 AATCTTGACACGCAACATTAACCTCACTCAATTCGCTATGCAACAGATGCGCAGGCGC 2768
 QY 2607 -----TAGCTAGCTCAAAACATAGCGCAG-----TGGCGCTTCAATTAGAGAGCGA 2653
 DB 2769 GCMAACCGGCACTGGGCAATGCGCGCGCGCGCTTCCGCGCTTCTTATTCGT 2828
 QY 2654 AACAGCGCAACATGCGCAACATGCTTTCAACATGCAACATGCAATGATTAATTGAG 2713
 DB 2829 TAGCGCGCAACTTGGAGAAATCCGCTTGAACGCTGACGCTAAGCGCAATTTGAA 2888
 QY 2714 TGGGCAAGCAATTCATTTACTTCACTTTATTTGGCTTAAAGCAATTTAA 2773
 DB 2889 CGGTACGAGAACATTCGCTTATGTCGMACTTGGGCTCCGAGCGCAATTTGAA 2948
 QY 2774 ATTATCCATAGCGTGAAGGCAATTCATTAATTTCTTCCGACACAGCGCAAGAC 2833
 DB 2949 GCTGGCGAAAGTTCCGAGGCACTTACCTTGGCTGCAAAATCCGCAACGAGC 3008
 QY 2834 CGAAACCTTGAGCAATTAATTTGCTGAAGCAAAATTAACATGATTAACGCTTATCAATTA 2893
 DB 3009 CGTAAGCTCGAGCAATTTGACGCTAGTGAAGAAAGAAACACACACGCGCTGCGAAA 3068
 QY 2894 GCTCAATTTACTTTAAGAAATGACACGCTTATGAGGAGCTTACCTTAAATTAAGT 2953
 DB 3069 TCTTAATTCACCTCGCAAAACGACACGCTGATCCGCGCATGCGCTTATCGCTTAC 3128
 QY 2954 GAAGATGATGCGGAATTCGCTGATCAATCAATTAAGCAAGCAAGATTTGCAATTA 3013
 DB 3129 CCGCAAGAGCGGCGATTCGCTGATTAATTCGCTGCAAAAGCAAGAGCTTTCCGCA 3188
 QY 3014 TTATAGTAGGAGCAAGCAAGCAACATTAAGCAACCAATTAAGCAAGCTGAC 3073
 DB 3189 ACTGCGAAGCGGAGAAACAGAGCGCGCTTGAACGCAAAACAGGCAAACTTGGCGC 3248
 QY 3074 TAAACACAAACAGCTGAGCAAAATGCGCTTGAAGAAAGCAAGCAAGCGCTTCC 3133
 DB 3249 CAACCAACAGCGGAAAGAAACAGCGCGCAAGCTTGAACGCTTATTCGCGCGCGC 3308
 QY 3134 TGATACCTGCTGATCAAGCTGTTAAACGCTTGAAGCAACCAATTAAGCTGAC 3193
 DB 3309 CAATGCCACCGAAAGCGCAAGAGCTGCTGCAACGCGCGCGCAAGCGGCGGAAAA 3368
 QY 3194 TGCTGAACACAAAAAGTAAGCAAAACAAAAAGATGCGCTCAAAAGAGAGTGT 3253
 DB 3369 TGCGGCAATTATGAGCGGAGGAAAGAAAGAGGCTGAGGAGGATTAAGACACCGC 3428
 QY 3254 TTCTGATCCCTGCTTATCAAGCTGTTCCATTTAGAGCGCACTTGAAGTTATGA 3313
 DB 3429 CTTGGCGAAACAGCGGAAAGCGAAACCGCGCGGCTTACACGCGCTTCCCGCGCGC 3488
 QY 3314 TGCGCCACAGCAATCGAAAAAGATGCTTGAAGCAAGAAAGCGGAAAGCAACGCA 3373
 DB 3489 CCGGCGCGCGGAGATTTGGCGCAACGCGCAAGCGCAACCGCAACCGCAAGCGG-- 3546
 QY 3374 ACAAAGACTTGATCAGCGCTTATTCAAAATGATGCTTATGAAATATCTGCAACGT 3433

DB 3547 -----GACCTATCAGCGCTTATGCAATGCGCTTGAAGTAATTTCCGCCAGCT 3589
 QY 3434 AATATGATCTTCTGTTCAAGATGAATTAATGCTTTTGTAGATCAAGCAATTC 3493
 DB 3600 CAACAGGCTTTTCCGCTTACAGAGCAATTTGACCGCGCTTTTCCGAAAGCGCGCA 3659
 QY 3494 TCCGCTGAGCAAAATATGCAAGATTAAGCAAGCTTATGATGCTTCCGCTC 3553
 DB 3660 GCGCTTTGAGCAAGCGGATCCGGGACACAAACATTAACGCTTGCAGATTTCCGCGC 3719
 QY 3554 TTATAGACGCAAAAGCAATTAATGCTTAAATGCGGCTGCAAAACCTTGAATG 3613
 DB 3720 CTACCGCAAC--AAACGCACTGCGCCAAATGCTTATGCAAGAAACCTTGGCGCG 3776
 QY 3614 AGCAATGCGGCGAGTTTCTGCAATGCGCTTCAATTAATGCTTATGATGAACGTTAA 3673
 DB 3777 GCGCTGCGATCTCTTTTTCGCAACCGGACCGGAAACCTTTCAGACGCGCATGCG 3836
 QY 3674 AATCAACGCAACATTAACGATGATGCGGCTTTCGCAATTAATGCGGCGATTTACA 3733
 DB 3837 CAATGCGCAACGCTTGCACAGGCTGCTTTCGCGCAATTAACGCGCATGCGGCTTGA 3896
 QY 3734 ATTGCTGTAACGCTGGAACGGGAATCACTGCGAGTAATGCTGTAAGCAAAAGCG 3793
 DB 3897 CATGCGCATGAGCGGCGCGGCTTTATGAGCGGAGCCCTTCAAGCGCATGAGG 3956
 QY 3794 AAAATTCATGCAAAAGCATTAATTAATGCGCTGAATGCAATTAATGCTTCCGTTAG 3853
 DB 3957 CAATTCGCGCGCGGCTGCTGATTAAGCAATTAATGCGCTGAATGCAATTAATGCTT 4016
 QY 3854 GCMAATGCGCAATTAATTAATTAATGAGTAATGCTTATTAATTAAGCAATTAATTA 3913
 DB 4017 CGGATTCGCAACGCAACCGCAATGCGCGCAACGCTTATTCGCAAAAGCGGATTA 4076
 QY 3914 TCAATCTAGAGATGAGATGAAGCAAGCTTACGCTTGAATTAATGCTTATTAATGCT 3973
 DB 4077 CGGATGCAAAAGCTCAATGCAATGCGCAACCGCGCTTGAATTAATGCTTATTAATGCT 4136
 QY 3974 CATGAGTTGATTAATTAATTAATTAATGCAATTAATTAATGCTTATTAATGCTTAT 4033
 DB 4137 CATTAAGCAATTAATTAATTAATTAATGCAATTAATTAATGCTTATTAATGCTTAT 4196
 QY 4034 CGTCAATTAATTAATTAATTAATTAATGCAATTAATTAATGCTTATTAATGCTTAT 4093
 DB 4197 CCGTCTTATGCAATGCGCTTCCGCAAGGCTGCAAGCGCGCTTATTAATGCTTATTAAT 4256
 QY 4094 GCAACACATTTGAGATTTATGCGCAAAAGAGGGAATTAAGGAGCAATTTTACA 4153
 DB 4257 GCGGAGGATTTGCGCAAAACCGGAGTGCGGATGCGGCTTAATGCGCAATTTTACA 4316
 QY 4154 TTTCGCAATTTCCGCTTATTAATTAATTAATGCAATTAATTAATGCTTATTAATGCTTAT 4213
 DB 4317 TTTCAGGCTTCCCTCAGCGCTGCGCGCGCAAGGCGGCAATTTGAAAGGAGGAGG 4376
 QY 4214 TGTGGCGTGAATTTGCGCTATGCTTGG 4241
 DB 4377 CCGGCGCATCAATTAATTAATTAATGCTTATGCTTGG 4404

RESULT 13

ACN43973

ID ACN40973 standard; DNA: 4404 BP.

XX AC ACN40973:

XX 27-OCT-2003 (revised)

DT 19-0UN-2003 (first entry)

XX DE Prokaryotic essential gene #22610.

XX KM Antisense; ds; prokaryotic essential gene; cell proliferation;

KM drug design; gene.

Db 1215 GATACCTACCAAGCAATCAACCAAGGCGGGGCTGTGATTTGAGGGTAATTTC 1274
Qy 1244 AGTATCTCCAAATTTCAACAACTTGGCAAGAGCTGGCATTCATGTAGTAAATAG 1303
Db 1275 GGTCTGGCTAAACCAACGAAACGTGGCAAGGCGGGGCTTCATTCAGTATGCGAG 1334
Qy 1304 CACCGTACTGGAAGAAATGGCGTGGACATGATCGACTTTCAAATTTGTTAAAG 1363
Db 1335 TACCGTACTGGAAGAAATGGCGTGGACATGATCGACTTTCAAATTTGTTAAAG 1394
Qy 1364 AACATTCAGCTTCAGGCAAGGGAATTAAGTTTCATGATGAGGATGAGGATGTA 1423
Db 1395 CAGCGTGTGCTTCAAGCAAAAGGGAATTAAGTTTCATGATGAGGATGAGGATGTA 1454
Qy 1424 AGTATTTGGAGCAGGCAAGGCAATCAAGGCAAAAGCTTTAGTAAATTTGG 1483
Db 1455 AGTCATCTTGAATCAGGCGGCGATCAAGGCAAAAGCTTTAGTAAATTTGG 1514
Qy 1484 CTGTGTTAGGCGCAAGGCAAGCTGTTCATTAACGATTAACATTTGATACGATTA 1543
Db 1515 CTGTGTTAGGCGCAAGGCAAGCTGTTCATTAACGATTAACATTTGATACGATTA 1574
Qy 1544 ATTTTATTTGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1603
Db 1575 ACTCATTTGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1634
Qy 1604 AGCTATCCAAATTCAGCAAGGCGCAATTTGTGATCAATTAACATTAACATTAAC 1663
Db 1635 CCGCATTCAAATTCAGCAAGGCGCAATTTGTGATCAATTAACATTAACATTAAC 1694
Qy 1664 TAAATGATCTATTTACGGAACGAAGCAATTTGTCTACCTTAATGAAATTAATTA 1723
Db 1695 CACGTTACATTTACGGAATTAACGGAATTAACGGAATTAACGGAATTAACGGA 1748
Qy 1724 ACTGATTTACGAAGAAATTTGCTTCAACGCTTGTGTGTGTGTGTGTGTGTGTGT 1783
Db 1749 CTGGAATGCAAAAGAAATTTGCTTCAACGCTTGTGTGTGTGTGTGTGTGTGTGT 1808
Qy 1784 ACACATTTGGCGCATTTAACTTATTTAAACCAACGGAATGATGATGATGATGAT 1843
Db 1809 AACGAAGGCGGCTCATTTGAATTAACCAACGGAATGATGATGATGATGATGAT 1868
Qy 1844 TTCAAGTGTACAAATTTAAAGGCGATTTACCAACAAAGGTAATTTTTCAG 1903
Db 1869 TTCCGCGGAAACAAATTTAAAGGCGATTTACCAACAAAGGTAATTTTTCAG 1928
Qy 1904 CGGTAGACGACACGCGCATTAACATTTAAATTAACGTTGTGTGTGTGTGTGTGT 1963
Db 1929 CGGAGACGACACGCGCATTAACATTTAAAGGCGGTTGTGTGTGTGTGTGTGT 1988
Qy 1964 TATACCAAGGGAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2023
Db 1989 TATCCCAAGGAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2048
Qy 2024 AAACCTTCAAATTTAAAGGCGAATTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2083
Db 2049 AAACCTTCAAATTTAAAGGCGAATTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2108
Qy 2084 AAATTTGACAGTCAGCAATTAAGCAATTTGCAATTTGTGTGTGTGTGTGTGTGT 2143
Db 2109 CGATTTGCAATTTAAAGCAATTTGCAATTTGTGTGTGTGTGTGTGTGTGTGTGT 2168
Qy 2144 TACATTTGCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2203
Db 2169 CACATCTGTACAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2228
Qy 2204 CGATCAAAAGTATTAATTTATACCAAAACAAATTCATGCTTATTAATTTTAC 2263
Db 2223 CGAGATTAAGTATTTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2288
Qy 2264 TGAATATGCAAGCGCATTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 2323
Db 2289 CGATCAAGCTCATTTAAATTTCAAGGAATTTGCAACCTCAACGCGCAATTTAGTACG 2348
Qy 2324 AAATCAGCA----- 2335
Db 2349 CGAGACACGCACTTAACGTTTACGCGCAACGCCAACCAAAAGCAACCTGAGCTGT 2408
Qy 2336 ----- 2335
Db 2409 GGGCAATGCCCAAGCAATTTAATCAAGCAATTTAAACGCAACATCGCTTCGA 2468
Qy 2336 -----ATTACATTAAGCAACATTTGCAACCAATTTAGCAATTTGCA 2386
Db 2469 CAATGCTTCAATTTAATTAAGCAACATTTGCAACCAATTTGCAACATTTGCA 2528
Qy 2387 CAATCAATCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2446
Db 2529 CAACCTTAAGCAAGTAAAGCAATTTGCAACATTTGCAACATTTGCAACATTTGCA 2588
Qy 2447 AGCTCAATTTCTTTAAACAGCAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2506
Db 2589 GCGATTTCAATTTGAAACAGCGCTTTACCGGAATTAACGCGCGCGCAAGGATAC 2648
Qy 2507 AACGTAAGTTGGAATTAAGCAATTTGCAACATTTGCAACATTTGCAACATTTGCA 2566
Db 2649 GCGATTTCAATTTAAGCAAGCAATTTGCAACATTTGCAACATTTGCAACATTTGCA 2708
Qy 2567 AACGTAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2606
Db 2709 AACCTTGAACAGCAACATTTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2768
Qy 2607 -----TCACTAGCTCAACATTAAGCAAG-----TCCGCTTATTAAGCA 2653
Db 2769 GCAAAACGCGAGTGCAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2828
Qy 2654 AACAGCGCAATTCGCGCAAGCAATTCGTTCAACATTTGCAACATTTGCAACATTTGCA 2713
Db 2829 TACCGCGCAATTCGCGCAAGCAATTCGTTCAACATTTGCAACATTTGCAACATTTGCA 2888
Qy 2714 TGGCAAGGCAATTCGCAATTTTCACTTCACTTTTGTGTGTGTGTGTGTGTGTGTGTGT 2773
Db 2889 CGGTCAAGCAATTTGCGCTTATGCGCAACCTTGTGTGTGTGTGTGTGTGTGTGTGT 2948
Qy 2774 ATTATGATACGCTGAGGCGATTTAATTTATTTGTGTGTGTGTGTGTGTGTGTGTGT 2833
Db 2849 GCTGCGGAAGTTCCAGAGCACTTACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3008
Qy 2834 CGAAACCTTGAAGCAATTTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2893
Db 3009 CGTATGCTCGAGCAATTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3068
Qy 2894 GCTCAATTTTACTTTAAGAAATGACACGTTGATGACGTTGATGATGATGATGATGAT 2953
Db 3069 TCTTATTTTCACTTGAAGAAAGCAATTTGATGATGATGATGATGATGATGATGATGAT 3128
Qy 2954 GAAATTAAGGCAATTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3013
Db 3129 CCGCAAGAGCGCAATTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3188
Qy 3014 TTTATTAAGAGAGCAAGCAAGCAATTTGAAAGCAAGTAAAGTAAAGTAAAGTAAAG 3073
Db 3189 ACTGCGCAAGGCGGAAGCAAGGCGCGCTTGAACGCAAGGCAAGGCAAGGCAAGGCA 3248
Qy 3074 TAAACCAACAGGTGAGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGG 3133
Db 3249 CAAGCAAGGCGGAAGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGG 3308
Qy 3134 TGAATCCCTGCTGATCAAGGCGCTGTAAAGCAATTTGAAAGCAAGTAAAGTAAAGTAA 3193
Db 3309 CAATGCGCAAGGCGGAAGCAAGGCGCGCTTGAACGCAAGGCAAGGCAAGGCAAGG 3368
Qy 3194 TGCCTGAAGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 3253
Db 3369 TGCCTGATTAAGGCGGAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 3428


```

QY 3254 TTCTGATCCCTGCTGATCAAGCCTGTCATAGAGACCGCATTTGAGTTATGTA 3313
Db 3429 CTTGGCGAAGACGGCGAAGCGAAGAACCGCGCTTACACCGCTTCCCGCGCGCG 3488
QY 3314 TGCCCCAGCAATCGAAGAAAAGATGCTGACTCAAGAAAGAGCGAAGAAACAGCGCA 3373
Db 3489 CCGCGCGCGCGGAGATTTGCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGG 3546
QY 3374 ACAAAGAGCTTGATGACCGCTTATTCATTAAGGCTTATCAAGATTAATCGCAACGT 3433
Db 3547 -----GACCTGATCAAGCGGTATGCTCAATGACGCTTATGAGTAAATTTCCGCGCGCT 3599
QY 3434 AATAGTATGCTTCTGCTGATCAAGTATGATGCTGCTTTTGTAGATCAAGACATC 3493
Db 3600 CAACAGCGTTTCCGCTTACAGACCAATGACCGGCTGTTGCGGAAAGCGCGCA 3659
QY 3494 TGCCGCTGCAAAATGCGACAGATTAAGACGCTTATGCTGATGCTGCTGCTG 3553
Db 3660 CGCGCTTGAAGACCGGCAATCGGAGACCAACACCTACCGCTGCAAGATTTCCGCGC 3719
QY 3554 TTATCGACGCAAGAAACGAACTTACGTAATTTGGGCTGCAAAAAGCTTACGATG 3613
Db 3720 CTACCGCCAAAC---AAACCGACCTGCGCAATGCTGACGAAACCTGCGACCG 3776
QY 3614 ACGAATTTGGGCGAGTTTCTGCAATGACGCTTCAATTAATCCTTTGATGAACAGTTAA 3673
Db 3777 GCGCGTGGCATCTCTGTTTTCGCAACCGGACCGGAAACCTTTCAGACGCGCATCG 3836
QY 3674 AAATCACCGGACATTAACGATGATGCGGCTTTTCCCAATATCAATGGGCGCATTTA 3733
Db 3837 CAACCTGGACCGCTTGGCCACGCTCCGCTTTTGGGCAATACGGGCAATCGCGAGTCA 3896
QY 3734 ATTGCTGTAACGCTGGGACGCGGATGAGTGGGTAATGGCTGAAGAAACAAACCG 3793
Db 3897 CATGGCATCAAGCGGCGCGCGGCTTATGAGCGGCAAGCTTTCAGACGCGCATCAG 3956
QY 3794 AAAAATTCATCGAAGAAAGCATTAATTAATGAGCTGTAAGTATCAAGTTCCGTTAG 3853
Db 3957 CAATATCCCGCGCGCGCTGCTGATTAAGGCAATTCAGGCAAGTACCGGCGAGGTTCCG 4016
QY 3854 GCAATTTGGGCACTGAGCTTATTTTGAAGTATGCTTATTTATGAAGTGAATTA 3913
Db 4017 CGGATTCGCGCATCGAAGCGGCAATCGGCGCAAGCGCTTATTTGTCGCAAAAAGCGATTA 4076
QY 3914 TCATCTGAGGAGGAGGAGGAGGAAAGCGCTTACGCTTCAATTAATGCTTAATGCTG 3973
Db 4077 CCGATGAGAAAGCGTAATATCGCAACCGCGGCGCTTTCATTAACCGCTTACCGCGCG 4136
QY 3974 CATTCGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4033
Db 4137 CATTAAGGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4196
QY 4034 GGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4093
Db 4197 CCGTCTCTATACCGAGCGCTTCCGCGCAAGTTCGAAACCGCGCTCAATTAATTAAT 4256
QY 4094 GCAACCAACATTTGAGAGCTTATTTGAGGAAAGAAAGTGGGATTAAGGAGGAAATTA 4153
Db 4257 GCGCGAGAGATTTGGCAAAACCGCGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGG 4316
QY 4154 TTTCGAATTTCCGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4213
Db 4317 TTTCAGCGCTGCTCCAGCGTGGCGCGCGCAAGGCGGCAATGGAAGCGGACAG 4376
QY 4214 TGTGGGCGTGAATTTGGGCTTGTGTTGG 4241
Db 4377 CGCGGCGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4404

```

RESULT 14
AAZ38918
ID AAZ38918 standard; DNA; 4374 BP.
XX

```

AC AAZ38918;
XX
DT 21-FEB-2000 (first entry)
XX
DE Neisseria meningitidis strain H44/76 BASB006 nucleotide sequence.
XX
KM Neisseria meningitidis; BASB006; diagnosis; bacterial; infection;
KM vaccine; antibiotic; upper respiratory tract infection; meningitis;
KM invasive bacterial disease; bacteraemia; screening; antibacterial; ss.
XX
OS Neisseria meningitidis.
XX
FH Key Location/Qualifiers
FT CDS 1..4374
FT FT /*tag= a
FT FT /product= "BASB006"
XX
PN W0995873..A2.
XX
PD 04-NOV-1999.
XX
PF 20-APR-1999; 99WO-EP002766.
XX
PR 24-APR-1998; 98GB-00008866.
XX
PA (SMIX ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PE Thomard J;
XX
DR WPI; 2000-052810/04.
XX
DR P-PSDB; AAT56622.
XX
PT Novel polynucleotides and polypeptides from Neisseria meningitidis used to
PT prepare vaccines against bacterial infections.
XX
PS Claim 11, Page 91-93; 103pp; English.
XX
XX
CC The present sequence encodes BASB006 isolated from Neisseria meningitidis
CC strain H44/76. BASB006 polynucleotides and polypeptides may be employed
CC as research reagents and material for the discovery of treatments and
CC diagnostics for diseases, particularly human diseases. They can be used
CC for diagnosis of disease, staging of disease, or determining response of
CC an infectious organism to drugs. The polynucleotides may be used as a
CC source for hybridisation probes, and for screening of genetic mutations,
CC serotype, organism or strain identification, identification of mutation
CC in BASB003 sequences, and as components of arrays which are useful for
CC diagnostic and prognostic purposes. The polypeptides can be used to
CC produce antibodies. The polypeptides can also be used in vaccine
CC formulations, and to identify agonists and antagonists. The polypeptides,
CC antibodies, agonists and antagonists (which are bacteriostatic) are used
CC for the treatment and prevention of diseases such as upper respiratory
CC tract infection, invasive bacterial diseases such as bacteraemia and
CC meningitis, and for the development and screening of antibacterial drugs.
CC They are also used in the prevention of adhesion of bacteria to
CC eukaryotic matrix proteins on in-dwelling devices, or to extracellular
CC proteins on wounds, and to prevent tissue damage and/or block the normal
CC progression of pathogenesis in infections initiated other than by the
CC implantation of in-dwelling devices or by other surgical techniques
XX
SQ Sequence 4374 BP; 1269 A; 1158 C; 1066 G; 881 T; 0 U; 0 Other;

```

Query Match 29.0%; Score 1252.6; DB 3; Length 4374;
Best Local Similarity 58.8%; Pred. No. 5.3e-282;
Matches 2567; Conservative 0; Mismatches 1534; Indels 264; Gaps 13;

```

QY 81 CTTAATTTTAAACGCTGTCATTTCAATTAAGTATGCAAGCGTGGCTGCTAC 140
Db 73 CCGCTTACTTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 132
QY 141 ACTAATTTTGGAGTATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 200
Db 133 ACTAATTTTGGAGTATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 192

```

QY 201 ACAGTGGGGCTCAAAATATTAGTTATTAACAACAGGCGAATTAGTTGGCAATCA 260
 DB 193 GCAATGGGGCGAAGAAATTTAGTTTACCAACAAAAGGGAAGTTGGTGGCAATCA 252
 QY 261 ATGACAAAAGCCCGATGATTTTCTGTAGTGTCACTAAGCGGTGGACCTTG 320
 DB 253 ATGACAAAAGCCCGATGATTTTCTGTAGTGTGGGTGGCGGTGGCGCAATG 312
 QY 321 GTTGAATATCAATATTTAGTGGACATGAAGTATGATATGATGATTT 380
 DB 313 GTGGGCGATCAATATTTAGTGGACATGAAGTATGATATGATGATTT 372
 QY 381 GTGGGCGATCAATATTTAGTGGACATGAAGTATGATATGATGATTT 440
 DB 373 GTGGGCGATCAATATTTAGTGGACATGAAGTATGATATGATGATTT 432
 QY 441 AACTCAAAAAG--ATATTTACCTTTATGGAAGATTCATATGATGATTT 497
 DB 433 AATATCAAGGAGGAGCAATGAAGTATGATATGATGATTT 492
 QY 498 CATTAATGTTACAGAGGAGGAGTATGATATGATGATTT 557
 DB 493 CATTAATGTTACAGAGGAGTATGATATGATGATTT 552
 QY 558 TATTCAGATGAGCAAAATATCAAGAGTGTGCTATGCTGTGAGCGGAGTTTG 617
 DB 553 TATTCAGATGAGCAAAATATCAAGAGTGTGCTATGCTGTGAGCGGAGTTTG 612
 QY 618 GGAATGATCAAGCAAGG--GACCAAGTGTGCTGTGATTT 659
 DB 613 GCAATGATCAAGCAAGG--GACCAAGTGTGCTGTGATTT 672
 QY 660 CATTAATGTTACAGAGGAGTATGATATGATGATTT 719
 DB 673 CATTAATGTTACAGAGGAGTATGATATGATGATTT 712
 QY 720 TTGGGAGCGAGTGTGTAAGCGGAGATATGATGATTT 779
 DB 733 TTGGGAGCGAGTGTGTAAGCGGAGATATGATGATTT 792
 QY 780 GGGGACAGTGTGTCCGATGTTATTTATGATGATTT 839
 DB 793 GGGGACAGTGTGTCCGATGTTATTTATGATGATTT 852
 QY 840 GGGGACAGTGTGTCCGATGTTATTTATGATGATTT 899
 DB 853 GGGGACAGTGTGTCCGATGTTATTTATGATGATTT 912
 QY 900 AATCTTATTT--TGATGAATTTTGAAGATTTATGATGATTT 956
 DB 913 AATCTTATTT--TGATGAATTTTGAAGATTTATGATGATTT 972
 QY 957 GCTGATATGAGTGTGATCAATTTAGTGAATTTATGATGATTT 1016
 DB 973 GCTGATATGAGTGTGATCAATTTAGTGAATTTATGATGATTT 1032
 QY 1017 AATCTGAGG--ATACATCAAAATTTAATTTAGTGAATTT 1058
 DB 1033 AATCTGAGG--ATACATCAAAATTTAATTTAGTGAATTT 1092
 QY 1059 ATGAGTTTACCTTTGAAGAGAGATTAAGTTCAATCTGATATGAGGAGT 1118
 DB 1093 GTTCTTTTACCTTTGAAGAGAGATTAAGTTCAATCTGATATGAGGAGT 1149
 QY 1119 ATTTATTTTCAAGTTTAAAGATGAGAGATTTATGATGATTT 1178
 DB 1150 ATTTATTTTCAAGTTTAAAGATGAGAGATTTATGATGATTT 1209
 QY 1179 TCAATATTTTCAAGTTTAAAGATGAGAGATTTATGATGATTT 1238
 DB 1210 TCAATATTTTCAAGTTTAAAGATGAGAGATTTATGATGATTT 1269
 QY 1239 TTTACATATTTCAAGTTTAAAGATGAGAGATTTATGATGATTT 1298

DB 1270 TTTAGGTTTGGCTGAGAAATTAAGAAACGTTGGAGGAGTGTGATGATGAA 1329
 QY 1299 AATGACCGCTTCTGTAAGTAAATGAGTGAACATGATGATGATTT 1358
 DB 1330 GACATGCTGCTGTAAGTAAATGAGTGAACATGATGATGATTT 1389
 QY 1359 AATGACCGCTTCTGTAAGTAAATGAGTGAACATGATGATGATTT 1418
 DB 1390 AATGACCGCTTCTGTAAGTAAATGAGTGAACATGATGATGATTT 1449
 QY 1419 GGTAAATGCTTTGAGACAGGAGTGAACATGATGATGATTT 1478
 DB 1450 GGTAAATGCTTTGAGACAGGAGTGAACATGATGATGATTT 1509
 QY 1479 ATGGCTTGTGAGGAGGAGTGAACATGATGATGATTT 1538
 DB 1510 ATGGCTTGTGAGGAGGAGTGAACATGATGATGATTT 1569
 QY 1539 GATTAATTTATTTGCGCTTTGAGGAGTGAACATGATGATTT 1598
 DB 1570 GATTAATTTATTTGCGCTTTGAGGAGTGAACATGATGATTT 1629
 QY 1599 TTTAAAGTATCCAAATATGAGGAGTGAACATGATGATTT 1658
 DB 1630 TTTAAAGTATCCAAATATGAGGAGTGAACATGATGATTT 1689
 QY 1659 GCGCTATGCTATTTACTGAGGAGTGAACATGATGATTT 1718
 DB 1690 GCGCTATGCTATTTACTGAGGAGTGAACATGATGATTT 1743
 QY 1719 AATTAATGCTATTTACAAATATGAGGAGTGAACATGATGATTT 1778
 DB 1744 AATTAATGCTATTTACAAATATGAGGAGTGAACATGATGATTT 1803
 QY 1779 AATTAATGCTATTTACAAATATGAGGAGTGAACATGATGATTT 1838
 DB 1804 AATTAATGCTATTTACAAATATGAGGAGTGAACATGATGATTT 1863
 QY 1839 GATCTTCAAGTGTGTAAGCGGAGATTTATGATGATTT 1898
 DB 1864 GATCTTCAAGTGTGTAAGCGGAGATTTATGATGATTT 1923
 QY 1899 TTTACGAGTGAAGGAGTGAACATGATGATTT 1958
 DB 1924 TTTACGAGTGAAGGAGTGAACATGATGATTT 1983
 QY 1959 GAGGATATCAAGGAGTGAACATGATGATTT 2018
 DB 1984 GAGGATATCAAGGAGTGAACATGATGATTT 2043
 QY 2019 GGTGAATCTTCAAAATTTAAGGAGTGAACATGATGATTT 2078
 DB 2044 GGTGAATCTTCAAAATTTAAGGAGTGAACATGATGATTT 2103
 QY 2079 GAGGAGTGAAGGAGTGAACATGATGATTT 2138
 DB 2104 GAGGAGTGAAGGAGTGAACATGATGATTT 2163
 QY 2139 GAAATATCATTTGAGGAGTGAACATGATGATTT 2198
 DB 2164 GAAATATCATTTGAGGAGTGAACATGATGATTT 2223
 QY 2199 TTTACGAGTGAAGGAGTGAACATGATGATTT 2258
 DB 2224 TTTACGAGTGAAGGAGTGAACATGATGATTT 2283
 QY 2259 TTTACGAGTGAAGGAGTGAACATGATGATTT 2318
 DB 2284 TTTACGAGTGAAGGAGTGAACATGATGATTT 2343
 QY 2319 TTTACGAGTGAAGGAGTGAACATGATGATTT 2378

Db 2344 GCAGATGGGATACAGCTTATACAGTCAGCCACAGCCGACCCGAAAGGCGACCTTAGC 2403
 QY 2379 ----- 2378
 Db 2404 CTGCTGGGCAATGCCCAGACATTTATACAGCCACATTAAAGCGACATCGGCT 2463
 QY 2379 -----CTT 2381
 Db 2464 TCGGCGAATGCTTCAATTAATCTAAGGACCAAGCCGTCAGAAAGCGAGTCAGGCTT 2523
 QY 2382 TCCGACATTCACCTGCAACGCTGATATGCAAACTTGAAGCGTAAATGTCATTTAG 2441
 Db 2524 TCCGCGAAGCTTAAGGCAACGTAAGCCATTCGCACTACAGGTAATGTCCTCAGCC 2583
 QY 2442 GATTACGCTCAATTTCTTTAAAAAGACCCATTTTTCGACCAATTCAGGAGACAA 2501
 Db 2584 GATAGGCAAGTATTCATTTGAAGAGAGCGCTTTACCGGCAAAATAGCGGCGGCAAG 2643
 QY 2502 GGCACACAGTGAAGTGGAAATGCAACTTGGACATGCTAGGATATCAATTCAG 2561
 Db 2644 GATACGCAATTCACCTTAAGACAGCCGATGAGACGTCGCTCAGGCAAGGAAATAGG 2733
 QY 2562 AATTACGCTAATTAACGATACGATCAAGTTAAATTCAGCTTAT 2606
 Db 2704 AATTAACTTGAACAGCCACATTAACATTCAGCTTATCGGCTATCGGCAAGATGCGCA 2763
 QY 2607 -----TACGCTAGCTCAAAATAGCCGACG-----TCGCGCT 2639
 Db 2764 GGGGCGCAACCGGAGTGCAGACAGTGCAGCGCGCGCGCTTCGCGCGCTTCGCGCGCT 2823
 QY 2640 TCATAGAGACGGAAACAGCCGACATCGGAGAAAGATGCTTTAAACATTCAGTAC 2699
 Db 2824 TCCCTATATTCGTTACCGGCACTTCGTAAGATCCGTTTCAACAGCTACGCTA 2883
 QY 2700 AATGTAAATTAAGTGGGCAAGGACATTCGATTAATTCATCTTATTTGGCTATTA 2759
 Db 2884 AAGCGCAATTTGAACGCTCAGGGAACATTCGCTTTATGTCGAACTTCGCGCTACCG 2943
 QY 2760 AGCGATAATTAATTAATTCATGACGCTGAGGCGATTCATATTAATTCGTTGCAAC 2819
 Db 2944 AGCGCAAAATTAAGCTGCGGAAAGTTCGAAAGGCACTTACCTTCGCGCTCAACAA 3003
 QY 2820 AAGCGCAAGAACCGGAAACCTTGAAGCAATTAATTCGTTGTAAGCAAGATATCA 2879
 Db 3004 ACGCGCAAGAACCTTGAAGCTTGAACATTTGACGTAAGTAAGGAAAGAAACAA 3063
 QY 2883 CCGTTATCAGATTAAGCTCAATTTACTTTAGAAATGAACAAGTTAGTCAAGTGCATTA 2939
 Db 3064 CCGCTGTCGAAATCTTTAATTTCACTTGAACAAAGACGTCATTCGCGCGCTG 3123
 QY 2940 CGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2999
 Db 3124 CGTTAACTCACTCCGAAAGAGCGGAGTTCCGCTGCAATATCCGCTCAAGAACAA 3183
 QY 3000 GAATTCGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3059
 Db 3184 GAGCTTTCCGAAACTGCGCAAGC-----AAGACCAAAA 3222
 QY 3060 GTTGAACGACCTGCTAAACAAACAGTGCAGCAAAAGTGCCTCAAGAAAGACGCG 3119
 Db 3222 CAGCGGAAAGAAACAAACGCGCAAGCTTGAACGCTGATTCGCGCGCGGAGATGCC 3282
 QY 3120 AGAGCAGCTTTCTGATACCTGCTGATCAAAAGCTGTTAAAGCATTAAGACCAA 3179
 Db 3283 GTCCAAAG---ACAGAAAGGCTTGCAGAACGCGCTCGAGGCGGAGGAAATGTC 3339
 QY 3180 CAAGTGAATGCTGCTGAACACAAATTAAGGCAAAACAAATTAAGTGCCTCA 3239
 Db 3340 GGCATTAATGAGCGGAGAAAGAAAGAAACGCTGTCAGGCAATTAAGCAACGCTTCG 3399
 QY 3240 AAAAGACAGCTGTTTCTGATCCCTGCTGATCAAAAGCTGTTGCTGATTAAGACGCA 3299
 Db 3400 GCGAAACAGCGGAGAGCGGAAACCGCGCTGACCAAGCTTCCTCCGCGCGCGCGC 3459

QY 3300 CTGAGGTTATGATGCCCAACAGCAATCGGAAAGATGCTTAGCTAAGAGAGAGCG 3359
 Db 3460 GCGCGCGGAGATTTGCGCGAATGGAACCCAAACGAGCCGCAAC----- 3506
 QY 3360 GAAAAACAGCAACAAAAAGACTTATAGCTGATTAATTAATTAATTAATTAATTA 3419
 Db 3507 -----GAGCGCGAAGCTGATCAGCGTTATAGCAATTAAGCTTGAAGTGA 3552
 QY 3420 TTAATGCAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3479
 Db 3553 TTTTCGCAAGCTCAACAGGCTTTGCGCTTGAAGAGAAATTAAGCCGATTTGCG 3612
 QY 3480 GATCAAGCAATCTCCGCTGACAAATATGCAACAGATTAAGACGCTATGATTC 3559
 Db 3613 GAGAGAGCGGCAACGCTTTGAGCAAGGAGCATCGGAGACCAACAACTACGCTTCG 3672
 QY 3540 GATGCTTCGCTCTTATAGCAAGCAAAAGCAATTAATTAATTAATTAATTAATTA 3599
 Db 3673 CAAGATTCGCGCCCTACCGCCAC---AAACGACCTTCGCAAAATGCTATGAGGAA 3729
 QY 3600 GCTTAGCTATGAGAGCAATGAGGAGCTTTCGCAATAGCCGTCAGATTAATTA 3659
 Db 3730 AACTTCGCAAGCGGCGCTGAGCAATCTGTTTTCGCAACCGGACCGAAACCTTC 3789
 QY 3660 GATGAACAGTTAAATCAACGCAATTAACGATGATGAGGTTTTCGCAATTA 3719
 Db 3790 GAGCAAGGATTCGCAACCTCGCAATCGGCTTCGCAAGCGGCGCTTTCGCGCAATACG 3849
 QY 3720 TGGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3779
 Db 3853 ATGCAAGATTTCAATTCGCAATCAAGCGGCGGCTTTTGAAGCGGACCTTTCA 3909
 QY 3780 GAGAGCAACCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3839
 Db 3910 GAGGCAATGAGGCAAAATCCGCGCGGCTGCTGATTAAGGCAATTCAGCAATAC 3969
 QY 3840 CAGTTCGCTTGAAGGCAATTCGCAATTCAGCTTATTTGAAGTAAATTCGCAATTA 3899
 Db 3970 CCGCGGCTTTCGCGGATTCGCAATTCAGCAATTCAGCTTATTTTCGCT 4029
 QY 3900 GAACTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3959
 Db 4030 CAAAAAGGAGATTCGCTACGAAACGCTCAATTAATTCAGCAATTCAGCTTATTA 4089
 QY 3960 CGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4019
 Db 4090 CGCTACCGCGGCAATTAAGGCAATTAATTAATTAATTAATTAATTAATTAATTA 4149
 QY 4150 AGCCCTAATTTGAGCCCTGCTTATACGATCGGCTTCGCGGAAAGTCCGAAACGCTC 4209
 Db 4080 AATTCAGGCTTGAACCAACCTTGAAGCTTATTCGCAAAAGAGAGGAGATTAAG 4139
 QY 4210 AATACCGCGTATGCTCAGATTTTCGCAAAACCGGAGTGCAGAAATGAGGAGTAAC 4269
 Db 4140 GCGAAATTTTAATTTCCAAATTCGCTTTATTCGCAAAATTCAGAGTTCACATC 4199
 QY 4270 GCGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4329
 Db 4200 GCGAAACAGCAAAATGAGGCGTGAATTAATTAATTAATTAATTAATTAATTAATTA 4244
 QY 4330 GAGGCGCAACAGCGGCGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4374

RESULT 15

AAZ12251

ID AAZ12251 standard; DNA; 4374 BP.

AC AAZ12251;

XX 08-OCT-1999 (first entry)

DT

XX Neisseria meningitidis complete ORF1 sequence.
 DE Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW Neisseria meningitidis; Neisseria infection; meningitis; septicemia; gonorrhea; ss.
 XX Neisseria meningitidis.
 OS Neisseria meningitidis.
 PN W03924578-A2.
 PD 20-MAY-1999.
 XX
 XX 03-OCT-1998; 98MO-1B001665.
 XX
 XX 06-NOV-1997; 97GB-00023516.
 XX 14-NOV-1997; 97GB-00024190.
 XX 18-NOV-1997; 97GB-00024386.
 XX 27-NOV-1997; 97GB-00025158.
 XX 10-DEC-1997; 97GB-00026147.
 XX 14-JAN-1998; 98GB-00000759.
 XX 01-SEP-1998; 98GB-00019016.
 XX (CHIR-) CHIRON SPA.
 XX
 XX Masignani V, Rapuoli R, Pizza M, Scarlato V, Grandi G;
 PI WPI; 1999-327407/27.
 DR P-PSDB; AAY38823.
 XX
 XX Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
 PT diagnosis, treatment and prevention of infection.
 XX
 PS Claim 9; Page 360-361; 524pp; English.
 XX
 XX Nucleotide sequences A211972-212358 represent open reading frames (ORFs)
 CC of Neisseria meningitidis and N. gonorrhoeae which encode antigenic
 CC proteins (see AAY38499-Y38944). The antigenic proteins, their fragments,
 CC their nucleic acids and antibodies are used for diagnosis, prevention (as
 CC vaccines) or treatment of Neisseria infections, such as meningitis,
 CC septicemia and gonorrhea. Both organisms are closely related. Fragments
 CC of the nucleic acids are useful as hybridisation probes and antisense
 CC reagents
 CC
 SC Sequence 4374 BP; 1266 A; 1163 C; 1066 G; 873 T; 0 U; 0 Other;
 Query Match 28.9%; Score 1245.4; DB 2; Length 4374;
 Best Local Similarity 58.8%; Pred. No. 3.2e-291;
 Matches 2565; Conservative 0; Mismatches 1536; Indels 264; Gaps 13;

QY 441 AACTACAAAAAG--ATATTTACATCTTTATGAGCGATTACCATATCCAGATTA 497
 DB 433 AATTATTAAGAGGACCTTAAGGCCATCTTTATGGGGCGATTATCATATGCCGTTT 492
 QY 498 CATTAATTCCTTACAGAGCGCTCCATTTGATATGATCTTGATATGATGGCTACT 557
 DB 493 CATTAATTTTCACAGATGCGAAGCCGTGTGAATGACAGTTATATGATGGCGGAAA 552
 QY 558 TATTCAGTATGACAAATATCCAGAAAGTTCGATTCGCTCGACCGCGTCTTGG 617
 DB 553 TATATCATATTAATATATATACCTGACCGGTTCGATTTGGGCGAGCGCAATATGG 612
 QY 618 CGAATGATCAGACAAAGGCG-----GACCAAGTTCGGGTGCATAT 659
 DB 613 CGATCATGATGAGATGAGCCCAATACCGGAAAGTTCAATCATATTTGCAATGGTAT 672
 QY 660 CATTAATCTGACAGCTGGCATATACAAATACCGGTGAGCGGTATATGATTTGGAT 719
 DB 673 TCTTGGCTCGTGTGTGGCATACCTTTGACAAATATGATGAGGTGGACAGTCAAC 732
 QY 720 TTGGAGGCGATGTTGCTTAAACGGGGAATATGCTCATTAACGATTGCGAGCTCAAG 779
 DB 733 TTAGGTATGAAAAATTAACATAGCCCATATGTTTTCACACAGAGGCTCATTT 792
 QY 780 GGGACAGTGGTTCCTCCAGTTTATTTATGATCTGAAAAACAAATGTTAAT 839
 DB 793 GGGACAGTGGCTCACCATATGTTTCTATGATGCCCAAAAGAGGTGTTAATAT 852
 QY 840 GGGATATTCGGGAAAGGCAACCTTTTGAAGCAAGAAATGGTTCAATTTGGTGGC 899
 DB 853 GGGATATTCGAAAGGCAACCTTTTGAAGCAAGAAATGGTTCAATTTGGTGGT 912
 QY 900 AATCTTATTT---TGATGAATTTTGAAGATTTTACATACATCATTTACACCGCA 956
 DB 913 AAGATATGTTCTATATGAAATCTTTGCGAGATACCATGATCTATACAGACA 972
 QY 957 GCTGTAAATGAGATGACAAATATGTAATATGATGTCAGGGTCTATATCTGAC 1016
 DB 973 CGTCAAAATGGAATATCTTTTACAGAGATATATAGGACAGAAAAATCAATGCGC 1032
 QY 1017 AATACGAG-----ATACATCGAATTAATTAATGATGATGAT 1058
 DB 1033 AATCATGAAACAAATTCCTGCTATATGATTAATAACAGAACCGTCAATGTTAAT 1092
 QY 1059 ATGATTAACCTTTGAAGAGAAAGATTAAGTCAATATCTAGATATAGCGACTAAT 1118
 DB 1093 GTTCTTATTCGAGACAGAGAAAGACCTGTTATATGCGCA---GGTGGTGCAC 1145
 QY 1119 AATTATCTCCAGGTTTAAACATGAGAAAGCTATATTTATGATCAAAAACAGCA 1178
 DB 1150 AGTTATCGACCAAGACTGATATATGAGAAATATTTCCTTTATGAGAAAGAAAGC 1209
 QY 1179 TCAATTAATTCGATCTGACATTAACAGAGGCGGAGTCTTATTTGAGGGTAT 1238
 DB 1210 GAATGATATCTTACACAGCAATCATATAGGTCTGAGAAATATATTTCAAGAGAT 1269
 QY 1239 TTACAGATCTCCAAATTTTACCAAACTTGCAAGAGCTGAGATACATGATAGTGA 1298
 DB 1270 TTTACGTCGCGCGGAAATATACAAATCTGCAAGGCGGGGTTCAATATAGTGA 1329
 QY 1299 AATGACCGTATCTTGGAAATGAAATGCGGTGAGACATGATGACTTTCTAATTTGGT 1358
 DB 1330 GACAGTACCGTATCTTGGAAATGAAATGCGGTGAGACATGATGACTTTCTAATTTGGT 1389
 QY 1359 AAGAGAACTTGGAGCTTCAAGCAAGGAGGAAATTAAGTTTCAATGAGGTGAGAT 1418
 DB 1390 AAGAGCAAGCTGAGAGTTCAGCAAGGAGGAAATCAAGGCTCATATGAGGTGGGAC 1449
 QY 1419 GGTAAAGTATTTTGAAGAGAGGACAGATCAAGGCAACAAACAGCTTTAGTGA 1478
 DB 1450 GGTAAAGTATTTTGAAGAGAGGACAGATCAAGGCAACAAACAGCTTTAGTGA 1509
 QY 1479 ATTGGCTGTTTGGGAGAGGAGCTGTCAATTAAGATGATTAACATTTGATACC 1538

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2004, 13:12:32 ; Search time 1688 Seconds

(without alignments)
11721.401 Million cell updates/sec

Title: US-10-645-655-1

Perfect score: 4319
Sequence: 1 TCATATGTCGTTACTAGT.....ACCTTTATCCATATAT 4319Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3017426 seqs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PCT_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4319	100.0	4319	US-09-839-996-1	Sequence 1, Appli
2	4319	100.0	4319	US-10-645-655-1	Sequence 1, Appli
3	4319	100.0	4319	US-10-080-505-1	Sequence 1, Appli
4	3180.8	73.6	4305	US-10-080-505-8	Sequence 8, Appli
5	3019.8	69.9	1830121	US-10-329-960-1	Sequence 1, Appli
6	3019.8	69.9	1830121	US-10-329-960-1	Sequence 1, Appli
7	2990.4	69.2	5245	US-10-080-505-12	Sequence 12, Appli
8	2970	68.8	4828	US-10-080-505-16	Sequence 16, Appli
9	2770.2	64.1	4822	US-10-080-505-14	Sequence 14, Appli
10	2709.6	62.7	4605	US-10-080-505-10	Sequence 10, Appli
11	1353.8	32.3	4350	US-10-282-122A-29527	Sequence 29527, A
12	1294.8	30.0	4404	US-10-282-122A-28843	Sequence 28843, A
13	202	4.7	5085	US-10-282-122A-22193	Sequence 22193, A
14	164.4	3.8	5322	US-10-282-122A-29633	Sequence 29633, A

15	155.2	2.6	4779	13	US-10-282-122A-29078	Sequence 29078, A
16	126.8	2.9	4113	15 <th>US-10-238-075-890</th> <th>Sequence 890, App</th>	US-10-238-075-890	Sequence 890, App
17	126.8	2.9	48012	15 <th>US-10-C85-953-251</th> <th>Sequence 251, App</th>	US-10-C85-953-251	Sequence 251, App
18	126.8	2.9	48254	15 <th>US-10-238-075-829</th> <th>Sequence 829, App</th>	US-10-238-075-829	Sequence 829, App
19	100.6	2.3	4128	15 <th>US-10-238-075-559</th> <th>Sequence 559, App</th>	US-10-238-075-559	Sequence 559, App
20	100.6	2.3	7654	15 <th>US-10-238-075-554</th> <th>Sequence 554, App</th>	US-10-238-075-554	Sequence 554, App
21	100.6	2.3	7654	15 <th>US-10-C85-953-15</th> <th>Sequence 15, Appl</th>	US-10-C85-953-15	Sequence 15, Appl
22	45.4	1.1	6378	15 <th>US-10-311-455-149</th> <th>Sequence 149, Appl</th>	US-10-311-455-149	Sequence 149, Appl
23	45.4	1.1	6283	15 <th>US-10-311-455-61</th> <th>Sequence 61, Appl</th>	US-10-311-455-61	Sequence 61, Appl
24	44.6	1.0	564	13 <th>US-10-027-632-50664</th> <th>Sequence 50664, A</th>	US-10-027-632-50664	Sequence 50664, A
25	44.6	1.0	564	13 <th>US-10-027-632-50665</th> <th>Sequence 50665, A</th>	US-10-027-632-50665	Sequence 50665, A
26	44.6	1.0	564	13 <th>US-10-027-632-66877</th> <th>Sequence 66877, A</th>	US-10-027-632-66877	Sequence 66877, A
27	44.6	1.0	564	13 <th>US-10-027-632-66878</th> <th>Sequence 66878, A</th>	US-10-027-632-66878	Sequence 66878, A
28	44.6	1.0	564	13 <th>US-10-027-632-66831</th> <th>Sequence 66831, A</th>	US-10-027-632-66831	Sequence 66831, A
29	44.6	1.0	564	13 <th>US-10-027-632-66832</th> <th>Sequence 66832, A</th>	US-10-027-632-66832	Sequence 66832, A
30	44.6	1.0	564	13 <th>US-10-027-632-70180</th> <th>Sequence 70180, A</th>	US-10-027-632-70180	Sequence 70180, A
31	44.6	1.0	564	13 <th>US-10-027-632-296122</th> <th>Sequence 296122, A</th>	US-10-027-632-296122	Sequence 296122, A
32	44.6	1.0	564	13 <th>US-10-027-632-296123</th> <th>Sequence 296123, A</th>	US-10-027-632-296123	Sequence 296123, A
33	44.6	1.0	564	13 <th>US-10-027-632-50664</th> <th>Sequence 50664, A</th>	US-10-027-632-50664	Sequence 50664, A
34	44.6	1.0	564	16 <th>US-10-027-632-50665</th> <th>Sequence 50665, A</th>	US-10-027-632-50665	Sequence 50665, A
35	44.6	1.0	564	16 <th>US-10-027-632-66877</th> <th>Sequence 66877, A</th>	US-10-027-632-66877	Sequence 66877, A
36	44.6	1.0	564	16 <th>US-10-027-632-66878</th> <th>Sequence 66878, A</th>	US-10-027-632-66878	Sequence 66878, A
37	44.6	1.0	564	16 <th>US-10-027-632-66831</th> <th>Sequence 66831, A</th>	US-10-027-632-66831	Sequence 66831, A
38	44.6	1.0	564	16 <th>US-10-027-632-66832</th> <th>Sequence 66832, A</th>	US-10-027-632-66832	Sequence 66832, A
39	44.6	1.0	564	16 <th>US-10-027-632-70179</th> <th>Sequence 70179, A</th>	US-10-027-632-70179	Sequence 70179, A
40	44.6	1.0	564	16 <th>US-10-027-632-70180</th> <th>Sequence 70180, A</th>	US-10-027-632-70180	Sequence 70180, A
41	44.6	1.0	564	16 <th>US-10-027-632-296122</th> <th>Sequence 296122, A</th>	US-10-027-632-296122	Sequence 296122, A
42	44.6	1.0	564	16 <th>US-10-027-632-296123</th> <th>Sequence 296123, A</th>	US-10-027-632-296123	Sequence 296123, A
43	44.6	1.0	564	13 <th>US-10-027-632-276009</th> <th>Sequence 276009, A</th>	US-10-027-632-276009	Sequence 276009, A
44	44.6	1.0	654	13 <th>US-10-027-632-276009</th> <th>Sequence 276009, A</th>	US-10-027-632-276009	Sequence 276009, A
45	44.6	1.0	654	16 <th>US-10-027-632-276009</th> <th>Sequence 276009, A</th>	US-10-027-632-276009	Sequence 276009, A

ALIGNMENTS

RESULT 1
US-09-839-996-1
Sequence 1, Application US/09839996
Publication No. US2003009010A1
GENERAL INFORMATION:
APPLICANT: St. Gene III, Joseph W.
TITLE OF INVENTION: Hemophilus Adherence and Penetration Protein
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSER: Fiehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/839,996
FILING DATE: 20-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/296,791
FILING DATE: 25-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Treacraft, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59941/NFT/RYS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1899
TELEFAX: (415) 398-3249
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4319 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: both

FEATURE:

NAME/KEY: CDS

LOCATION: 60..4241

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-B39-996-1

Query Match 100.0%; Score 4319; DB 10; Length 4319;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAATAGCTGTTAACTAGATATTTTAAATGCAAAATTTCTAAATTAATTAACCTTA 60
 DB 1 TCAATAGCTGTTAACTAGATATTTTAAATGCAAAATTTCTAAATTAATTAACCTTA 60
 QY 61 TGAATAAACTGTAATTTGCTGTAATTTTAAACGCTTGATTCATTAAGGATAGAT 120
 DB 61 TGAATAAACTGTAATTTGCTGTAATTTTAAACGCTTGATTCATTAAGGATAGAT 120
 QY 121 CGCAAGCGTGGCTGGTCACTTAATTTGGGATTGATTACCAATTAATTCGTGATTTG 180
 DB 121 CGCAAGCGTGGCTGGTCACTTAATTTGGGATTGATTACCAATTAATTCGTGATTTG 180
 QY 181 CCGAAGATTAAGGAAAGTTCAAGTTGGGCTGCAAAATTTTAAGCTTTAATAAACAAG 240
 DB 181 CCGAAGATTAAGGAAAGTTCAAGTTGGGCTGCAAAATTTTAAGCTTTAATAAACAAG 240
 QY 241 GCGAATAGTGGCACTCAATGACAAAGCCCGATGATGATTTTCTGTAAGTGCAC 300
 DB 241 GCGAATAGTGGCACTCAATGACAAAGCCCGATGATGATTTTCTGTAAGTGCAC 300
 QY 301 GTACGCGCTGGACGCTTGGTGAATAATATATTTGAGCGTGGCAATTAACGATG 360
 DB 301 GTACGCGCTGGACGCTTGGTGAATAATATATTTGAGCGTGGCAATTAACGATG 360
 QY 361 GATATACAGATGTTGATTTGGTGCAGAGGAAACAACCCGATCAATCGTTTACTT 420
 DB 361 GATATACAGATGTTGATTTGGTGCAGAGGAAACAACCCGATCAATCGTTTACTT 420
 QY 421 ATTAAGATTTAAACAAATTAATCTACAAATTAATTAATCTCTATAGAGACAT 480
 DB 421 ATTAAGATTTAAACAAATTAATCTACAAATTAATTAATCTCTATAGAGACAT 480
 QY 481 ACCATTAATCAAGATTAATTAATTTGTTACAGAGCGGCTCCAAATGATGATGACTTGA 540
 DB 481 ACCATTAATCAAGATTAATTAATTTGTTACAGAGCGGCTCCAAATGATGATGACTTGA 540
 QY 541 ATATGATGCGAGTACTTATTCAGATGAAACAAATATCCGAAATGATGATGATGAT 600
 DB 541 ATATGATGCGAGTACTTATTCAGATGAAACAAATATCCGAAATGATGATGATGAT 600
 QY 601 CTGAGCGCGAGTTTGGGAAATGATCAAGACAAAGGACCAAGTTCGCGGTGATC 660
 DB 601 CTGAGCGCGAGTTTGGGAAATGATCAAGACAAAGGACCAAGTTCGCGGTGATC 660
 QY 661 ATTATCTGACAGCTGGCAATACCAATCAAGCGTGGAGCAGATGATGATGATGAT 720
 DB 661 ATTATCTGACAGCTGGCAATACCAATCAAGCGTGGAGCAGATGATGATGATGAT 720
 QY 721 TGGGAGGCGAGTGTCTGTAAGCGGAAATATGTCATTCACGATTCGAGGTCAAGG 780
 DB 721 TGGGAGGCGAGTGTCTGTAAGCGGAAATATGTCATTCACGATTCGAGGTCAAGG 780
 QY 781 GGGACAGTGTCTCCGATGTTATATGATCTGTAAGGAAACAAATATGTTATTAATG 840
 DB 781 GGGACAGTGTCTCCGATGTTATATGATCTGTAAGGAAACAAATATGTTATTAATG 840
 QY 841 GGAATATTAAGGAAAGCAACCTTTTGAAGGCAAGAAATGGGTTTCAATGGTTCCCA 900

DB 841 GGAATATTAAGGAAAGCAACCTTTTGAAGGCAAGAAATGGGTTTCAATGGTTCCCA 900
 QY 901 AACTTATTTTGAAGAAATTTTGAAGATTTTACATACATCACTTAACCCGAGCTG 960
 DB 901 AACTTATTTTGAAGAAATTTTGAAGATTTTACATACATCACTTAACCCGAGCTG 960
 QY 961 GTATGAGTGTACACATTAAGTGAATGATGATGATGATGATGATGATGATGAT 1020
 DB 961 GTATGAGTGTACACATTAAGTGAATGATGATGATGATGATGATGATGATGAT 1020
 QY 1021 CAGAAATACATGCAATTAATTAATTTACCTTACCAATTAATTTACCTTACCAAT 1080
 DB 1021 CAGAAATACATGCAATTAATTAATTTACCTTACCAATTAATTTACCTTACCAAT 1080
 QY 1081 AGATTAAGTGTACCAATTAATTAATTTACCTTACCAATTAATTTACCTTACCAAT 1140
 DB 1081 AGATTAAGTGTACCAATTAATTAATTTACCTTACCAATTAATTTACCTTACCAAT 1140
 QY 1141 ATGGAATAACCTTAATTTATGATGATCAAAACAAAGGATCAATTAATTTACCTTACCAAT 1200
 DB 1141 ATGGAATAACCTTAATTTATGATGATCAAAACAAAGGATCAATTAATTTACCTTACCAAT 1200
 QY 1201 TTAACCAAGGCGGCTGCTTATTTTGAAGGATTTTACATGATTCGCAATTTCA 1260
 DB 1201 TTAACCAAGGCGGCTGCTTATTTTGAAGGATTTTACATGATTCGCAATTTCA 1260
 QY 1261 ACCAATCTTGGCAAGAGCTGGGATCAATGATGAAATGACACCTTACCTTGAAG 1320
 DB 1261 ACCAATCTTGGCAAGAGCTGGGATCAATGATGAAATGACACCTTACCTTGAAG 1320
 QY 1321 TAATGAGTGGAAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 DB 1321 TAATGAGTGGAAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 QY 1381 CCAAGGCGGAAATTAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1440
 DB 1381 CCAAGGCGGAAATTAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1440
 QY 1441 AGGCAAGATCAAGGCAAGGCAAGGCTTACGAAATTTGCTTGGTGGCTGAG 1500
 DB 1441 AGGCAAGATCAAGGCAAGGCAAGGCTTACGAAATTTGCTTGGTGGCTGAG 1500
 QY 1501 GGAATGCTTAATTAAGATGATTAACCAATTTGATGATGATGATGATGATGATGAT 1560
 DB 1501 GGAATGCTTAATTAAGATGATTAACCAATTTGATGATGATGATGATGATGATGAT 1560
 QY 1561 GTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 DB 1561 GTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 QY 1621 ACAGAGGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 DB 1621 ACAGAGGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 QY 1681 GGAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 DB 1681 GGAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 QY 1741 AATGCTGCTCAAGGCTTGGTGGGAAACGATTAATTAATTAATTAATTAATTAAT 1800
 DB 1741 AATGCTGCTCAAGGCTTGGTGGGAAACGATTAATTAATTAATTAATTAATTAAT 1800
 QY 1801 AACTTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860
 DB 1801 AACTTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860
 QY 1861 TAAAGGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920
 DB 1861 TAAAGGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920
 QY 1921 AAGCTTCAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1980

Db	1921	ACGCTTCAATCACTTTAAATTAAGTTGTGCGAAGAAATGGAATGATATCCACCAAGCGCAA	1980
Qy	1981	TTGTGTGGGATGACGATTTGATCAACCGTACATTTAAAGCTGAAACCTTCCAAATTAAG	2540
Db	1981	TTGTGTGGGATGACGATTTGATCAACCGTACATTTAAAGCTGAAACCTTCCAAATTAAG	2040
Qy	2041	GGGAAATGGCGGTGGTTCTCGCAATGTTCTTCCAAATGAGGGAAATGGAGCGTCA	2100
Db	2041	GGGAAATGGCGGTGGTTCTCGCAATGTTCTTCCAAATGAGGGAAATGGAGCGTCA	2100
Qy	2101	ATAATGCAAAATGCGACATTTGGGTGTGCGCAATCACAAAATATCAATTGCGCGTT	2160
Db	2101	ATAATGCAAAATGCGACATTTGGGTGTGCGCAATCACAAAATATCAATTGCGCGTT	2160
Qy	2161	CAGATTGGACGAGATTAAAGACTTGTCCAAAAGTGTATTAACCGATCAAAAGTTATTA	2220
Db	2161	CAGATTGGACGAGATTAAAGACTTGTCCAAAAGTGTATTAACCGATCAAAAGTTATTA	2220
Qy	2221	ATTCTATACCAAAAACACAAATCAATGCGCTCTATTATTAATCGATTAATGCAAGGGGA	2280
Db	2221	ATTCTATACCAAAAACACAAATCAATGCGCTCTATTATTAATCGATTAATGCAAGGGGA	2280
Qy	2281	ATGTTAAAGGTTTAGCAAACTTAATGSCATGTCACTTTAAACAATACAGCGCAATTA	2340
Db	2281	ATGTTAAAGGTTTAGCAAACTTAATGSCATGTCACTTTAAACAATACAGCGCAATTA	2340
Qy	2341	CATTAAAGCAACAATGCCACCCAAATAGGCAATTTGCACTTTCGACAAATTCACACTGAA	2400
Db	2341	CATTAAAGCAACAATGCCACCCAAATAGGCAATTTGCACTTTCGACAAATTCACACTGAA	2400
Qy	2401	CGGTGATTAATGCAAACTTGAACGGTATGTGCATTTAAACGATTTGACGTCAATTTCTT	2460
Db	2401	CGGTGATTAATGCAAACTTGAACGGTATGTGCATTTAAACGATTTGACGTCAATTTCTT	2460
Qy	2461	TAAAAAAGCATTCTTGTGACACAAATTCAGGAGACAAAGSCAACAGTACGCTTGG	2520
Db	2461	TAAAAAAGCATTCTTGTGACACAAATTCAGGAGACAAAGSCAACAGTACGCTTGG	2520
Qy	2521	AAAATGCACTTGTGACATGCTTACGATACATTAATTCGAAATTTAAGCTTAATATACA	2580
Db	2521	AAAATGCACTTGTGACATGCTTACGATACATTAATTCGAAATTTAAGCTTAATATACA	2580
Qy	2581	GTACGATCACTTAAATTCAGCTTATTCAGCTGACCTCAACAAATACGCAAGTGCCTT	2640
Db	2581	GTACGATCACTTAAATTCAGCTTATTCAGCTGACCTCAACAAATACGCAAGTGCCTT	2640
Qy	2641	CATTGAGACGGAACCAAGCCCAACATGCGAGACATGCTTTCACACATTAACAGTAA	2700
Db	2641	CATTGAGACGGAACCAAGCCCAACATGCGAGACATGCTTTCACACATTAACAGTAA	2700
Qy	2701	ATGATTAATTTAGTGGGCAAGACATTCGAAATTTACTCATCTTATTTGGCTATTA	2760
Db	2701	ATGATTAATTTAGTGGGCAAGACATTCGAAATTTACTCATCTTATTTGGCTATTA	2760
Qy	2761	GGCATTAATTTAAATTTATCCATGACGCTGAGGCGCATTAACATATATCTGTTGGAGAA	2820
Db	2761	GGCATTAATTTAAATTTATCCATGACGCTGAGGCGCATTAACATATATCTGTTGGAGAA	2820
Qy	2821	CAGGCAAGAAACCCGAAACCTTGAGCAATTTACTTTGGTGAAGCAAAATATATCAAC	2880
Db	2821	CAGGCAAGAAACCCGAAACCTTGAGCAATTTACTTTGGTGAAGCAAAATATATCAAC	2880
Qy	2881	CGTTATCAGATTAAGCTCAAACTTACTTTAGAAAATGACACGTTGATGACAGGTGCAATAC	2940
Db	2881	CGTTATCAGATTAAGCTCAAACTTACTTTAGAAAATGACACGTTGATGACAGGTGCAATAC	2940
Qy	2941	GTATTAATTTAGTGAAGATGATGCGCAATTTCCCTTGCATTAACCCAAATTAATAATAG	3000
Db	2941	GTATTAATTTAGTGAAGATGATGCGCAATTTCCCTTGCATTAACCCAAATTAATAATAG	3000
Qy	3001	AATTTGACAAATGTTTGTGTAAGACGAGACAGACAGACATTAACCAACCCAAACAG	3060
Db	3001	AATTTGACAAATGTTTGTGTAAGACGAGACAGACAGACATTAACCAACCCAAACAG	3060

QY	3061	TTGAACCGATCTGTAAACACAAACAGGCTGACCTCAAAAGTGGGCTCAAGAAAGCAGCGA	3120
Db	3061	TTGAACCGATCTGTAAACACAAACAGGCTGAGCCAAAGTGGGCTCAAGAAAGCAGCGA	3120
QY	3121	GAGCAGCGGTTCCTGATACCTCTGCTATCAAGCCTTTAAACGATTAGAACCAAC	3180
Db	3121	GAGCAGCGGTTCCTGATACCTCTGCTATCAAGCCTTTAAACGATTAGAACCAAC	3180
QY	3181	AACTGAACCTGACTGTGAACACAAAAAAGTAAAGCAAAAAACAAAAAAGTGGGCTAA	3240
Db	3181	AACTGAACCTGACTGTGAACACAAAAAAGTAAAGCAAAAAACAAAAAAGTGGGCTAA	3240
QY	3241	AAAGAGAGGTTCCTGATACCTCTGCTATCAAGCCTTTAAACGATTAGAACCAAC	3300
Db	3241	AAAGAGAGGTTCCTGATACCTCTGCTATCAAGCCTTTAAACGATTAGAACCAAC	3300
QY	3301	TTGAGTTATGATGCCCCACAGCAACCGGAAAGATGGCTGCTCAAGAAAGAGCG	3360
Db	3301	TTGAGTTATGATGCCCCACAGCAACCGGAAAGATGGCTGCTCAAGAAAGAGCG	3360
QY	3361	AAAAACAACGCAACAAAAAGACTTGAATCGACCGGTATTCAAATAGTGGCTTACGAT	3420
Db	3361	AAAAACAACGCAACAAAAAGACTTGAATCGACCGGTATTCAAATAGTGGCTTACGAT	3420
QY	3421	TATCTGCAACAGTAAATAGTATGCTTCTGTCAAGATGAATAGATCTCTTTTGAG	3480
Db	3421	TATCTGCAACAGTAAATAGTATGCTTCTGTCAAGATGAATAGATCTCTTTTGAG	3480
QY	3481	ATCAGACCAATCTGCGCGTGTGACAAATATCGCACAGGATTAAGACGCTATGATCG	3540
Db	3481	ATCAGACCAATCTGCGCGTGTGACAAATATCGCACAGGATTAAGACGCTATGATCG	3540
QY	3541	ATGCGTTCCGCTCTATCAGCAGCAAGAAACGAATCTTACGTCAATTGCGGTTCAAAAG	3600
Db	3541	ATGCGTTCCGCTCTATCAGCAGCAAGAAACGAATCTTACGTCAATTGCGGTTCAAAAG	3600
QY	3601	CCTTAGCTAATGACGAATGGGGCAGTTTTCTCGATAGCGCTTCAGATATACCTTG	3660
Db	3601	CCTTAGCTAATGACGAATGGGGCAGTTTTCTCGATAGCGCTTCAGATATACCTTG	3660
QY	3661	ATGAAACAGTAAATATCAGCGACATTAACGATGATGCGGCTTTTCCCAATATCAAT	3720
Db	3661	ATGAAACAGTAAATATCAGCGACATTAACGATGATGCGGCTTTTCCCAATATCAAT	3720
QY	3721	GGGGCGATTTACATTTGGTGTAAACGTGGAAACGGAAATCAATGACGATTAATGCTCG	3780
Db	3721	GGGGCGATTTACATTTGGTGTAAACGTGGAAACGGAAATCAATGACGATTAATGCTCG	3780
QY	3781	AAAGACAAAGCCGAAAAATTCATCGAAAGCGATTAATTAAGCGGTGAATGCAAGTTATC	3840
Db	3781	AAAGACAAAGCCGAAAAATTCATCGAAAGCGATTAATTAAGCGGTGAATGCAAGTTATC	3840
QY	3841	AGTTCGTTTGAAGCAATGGGGAATTCAGCCTTAATTTGAGATTAATGCAATTTTATG	3900
Db	3841	AGTTCGTTTGAAGCAATGGGGAATTCAGCCTTAATTTGAGATTAATGCAATTTTATG	3900
QY	3901	AACTGAAATTTATCAATCTTGAAGAAGTGAAGTGAAGAACGCTTGAAGCTTCATTAATC	3960
Db	3901	AACTGAAATTTATCAATCTTGAAGAAGTGAAGTGAAGAACGCTTGAAGCTTCATTAATC	3960
QY	3961	GCATTAAGTGGGCAATTCGATGATTAATCAATTTCTCGACAGATTAATATCAGCGTTA	4020
Db	3961	GCATTAAGTGGGCAATTCGATGATTAATCAATTTCTCGACAGATTAATATCAGCGTTA	4020
QY	4021	AGCCTTAATTTCTGCAATTAATGATGTTTAAACGTTAAGTAAACCAACGCTTAA	4080
Db	4021	AGCCTTAATTTCTGCAATTAATGATGTTTAAACGTTAAGTAAACCAACGCTTAA	4080
QY	4081	ATCTCAGGCTGTGCAACCAATTTGACGCTTAATGGCAAAAAAAGTGGGATTTAAAG	4140
Db	4081	ATCTCAGGCTGTGCAACCAATTTGACGCTTAATGGCAAAAAAAGTGGGATTTAAAG	4140

QY 4141 CAGAAATTACATTCGAAATTCGCTTTATCTCAAAATCTCAAGGTCACAACTCG 4200
 DB 4141 CAGAAATTACATTCGAAATTCGCTTTATCTCAAAATCTCAAGGTCACAACTCG 4200
 QY 4201 GCAACAGCAAAATGTGGCGGTGAAATTTGAGCTATCGTTGTAATAATCAATATTTT 4260
 DB 4201 GCAACAGCAAAATGTGGCGGTGAAATTTGAGCTATCGTTGTAATAATCAATATTTT 4260
 QY 4261 ATCGTTTATGATTAACAGGTGGTGCATGATGATCCCACTTTTATTCGAATAT 4319
 DB 4261 ATCGTTTATGATTAACAGGTGGTGCATGATGATCCCACTTTTATTCGAATAT 4319

RESULT 2

US-10-645-655-1
 ; Sequence 1, Application US/10645655
 ; Publication No. US20040063908A1
 ; GENERAL INFORMATION:

APPLICANT: St. Gene III, Joseph W.
 Palkow, Stanley

TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
 STREET: 4 Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: California
 COUNTRY: United States
 ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/645,655
 FILING DATE: 20-Aug-2003
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/296,791
 FILING DATE: 25-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: Treacartin, Richard F.
 REGISTRATION NUMBER: 31,801

TELEPHONE: (415) 781-1989
 TELEFAX: (415) 398-3249

REFERENCE/DOCKET NUMBER: A-59941/RET/RMS

INFORMATION FOR SEQ. ID NO. 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4319 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: both

FEATURE:

NAME/KEY: CDS
 LOCATION: 60..4241

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-645-655-1

Query Match 100.0%; Score 4319; DB 13; Length 4319;
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAATAGTCGTTTAACTAGTATTTTAAATGCAAAAATTAATTAATTAATTAATTAATTA 60
 DB 1 TCAATAGTCGTTTAACTAGTATTTTAAATGCAAAAATTAATTAATTAATTAATTAATTA 60
 QY 61 TGAATAAACTGTAATTTGCTTAATTTTAAACCGCTGCAATTCATGAGGATATAT 120
 DB 61 TGAATAAACTGTAATTTGCTTAATTTTAAACCGCTGCAATTCATGAGGATATAT 120

QY 121 CCAGAGGCTGGGCTGGTCAACTATTATTGGATTGATTACCAATATATGATTTTG 180
 DB 121 CCAGAGGCTGGGCTGGTCAACTATTATTGGATTGATTACCAATATATGATTTTG 180
 QY 181 CCGAGATTAAGGGAAGTTCAAGTGGGCTTAATTTAAGGTTTAAACAAAG 240
 DB 181 CCGAGATTAAGGGAAGTTCAAGTGGGCTTAATTTAAGGTTTAAACAAAG 240
 QY 241 GCGAATAGTTGGCAATCATGACAAAGAGCCCAATGATTTTCTGATGTCAC 300
 DB 241 GCGAATAGTTGGCAATCATGACAAAGAGCCCAATGATTTTCTGATGTCAC 300
 QY 301 GTTACGCGGTGGAGGCTGTTGAAATCATATATTTGAGCGTGGCAATACATGAG 360
 DB 301 GTTACGCGGTGGAGGCTGTTGAAATCATATATTTGAGCGTGGCAATACATGAG 360
 QY 361 GATATACGATGTGATTTTGGTGAAGGGAACAAACCCGATCAATGCTTTACTT 420
 DB 361 GATATACGATGTGATTTTGGTGAAGGGAACAAACCCGATCAATGCTTTACTT 420
 QY 421 ATTAAGATTGTAAACGAAATTAATCAAAAAGATTAATTAATCTTATGAGAGAT 480
 DB 421 ATTAAGATTGTAAACGAAATTAATCAAAAAGATTAATTAATCTTATGAGAGAT 480
 QY 481 ACCATATTCAGAGTTTACATTAATTTGTTACAGAGCGGCTCAATTTGATGACTTGA 540
 DB 481 ACCATATTCAGAGTTTACATTAATTTGTTACAGAGCGGCTCAATTTGATGACTTGA 540
 QY 541 AATGTAATGCGATCTTATGATGATTAATTAATTAATTAATTAATTAATTAATTA 500
 DB 541 AATGTAATGCGATCTTATGATGATTAATTAATTAATTAATTAATTAATTAATTA 500
 QY 541 AATGTAATGCGATCTTATGATGATTAATTAATTAATTAATTAATTAATTAATTA 500
 DB 541 AATGTAATGCGATCTTATGATGATTAATTAATTAATTAATTAATTAATTAATTA 500
 QY 601 CTGAGAGGAGTTTGGCGAATGATCAAGCAAAAGGCGAATGTCGATGATC 660
 DB 601 CTGAGAGGAGTTTGGCGAATGATCAAGCAAAAGGCGAATGTCGATGATC 660
 QY 661 ATTATCTGACGCTGCAATACCAATATGCTGAGAGGATGATGATTTGATTT 720
 DB 661 ATTATCTGACGCTGCAATACCAATATGCTGAGAGGATGATGATTTGATTT 720
 QY 721 TGGAGAGGAGTTTGGCGAATGATCAAGCAAAAGGCGAATGTCGATGATC 780
 DB 721 TGGAGAGGAGTTTGGCGAATGATCAAGCAAAAGGCGAATGTCGATGATC 780
 QY 781 GGGACAGTGGTCTCCGATGTTTATTAATGATGCTGAAAAAAGGTTAATTAATG 840
 DB 781 GGGACAGTGGTCTCCGATGTTTATTAATGATGCTGAAAAAAGGTTAATTAATG 840
 QY 841 GGAATTAAGGGAAGGCAACCTTTTGAAGCAAAAGGTTTGAATGCTTCCGA 900
 DB 841 GGAATTAAGGGAAGGCAACCTTTTGAAGCAAAAGGTTTGAATGCTTCCGA 900
 QY 901 AATCTTAATTTGAATTAATTTGAAAGATTTGAAAGATTTGAAAGATTTGAAAGATTT 960
 DB 901 AATCTTAATTTGAATTAATTTGAAAGATTTGAAAGATTTGAAAGATTTGAAAGATTT 960
 QY 961 GTATGAGATGATGATTAATTTGAAAGATTTGAAAGATTTGAAAGATTTGAAAGATTT 1020
 DB 961 GTATGAGATGATGATTAATTTGAAAGATTTGAAAGATTTGAAAGATTTGAAAGATTT 1020
 QY 1021 CAGGAATACATGCAAAATTAATTAATTTAGTTGCAAAATGAGTTTACCTTTGAAGAG 1080
 DB 1021 CAGGAATACATGCAAAATTAATTAATTTAGTTGCAAAATGAGTTTACCTTTGAAGAG 1080
 QY 1081 AGATTAATGATTAATTTGAAAGATTTGAAAGATTTGAAAGATTTGAAAGATTTGAAAG 1140
 DB 1081 AGATTAATGATTAATTTGAAAGATTTGAAAGATTTGAAAGATTTGAAAGATTTGAAAG 1140
 QY 1141 ATGAGAAAGCTAATTTTATGATGATCAAAAGGATCAATTAATTTGATGATCA 1200
 DB 1141 ATGAGAAAGCTAATTTTATGATGATCAAAAGGATCAATTAATTTGATGATCA 1200

1201 TTAACCAAGGCGGGGCTCTTATTTTGAAGGTAATTTACAGTATCTCCAAATCTTA 1260
 1201 TTAACCAAGGCGGGGCTCTTATTTTGAAGGTAATTTACAGTATCTCCAAATCTTA 1260
 1261 ACCAACTTGGCAAGAGCTGGCATACATGTAAGTAAATAGCAACGGTACTTGAAG 1320
 1261 ACCAACTTGGCAAGAGCTGGCATACATGTAAGTAAATAGCAACGGTACTTGAAG 1320
 1321 TAAATGGGTGGAACATGATGCACTTTCTAAATTTGGTAAAGGAACATTTGAGCGTCAAG 1380
 1321 TAAATGGGTGGAACATGATGCACTTTCTAAATTTGGTAAAGGAACATTTGAGCGTCAAG 1380
 1381 CCAAGGCGGAAATTAAGGTTGATGAGGAGGAGTGGTAAAGTCAATTTTGGAGAGC 1440
 1381 CCAAGGCGGAAATTAAGGTTGATGAGGAGGAGTGGTAAAGTCAATTTTGGAGAGC 1440
 1441 AGGCAAGCATCAAGGCAACAAACAGCTTTAGTAAATTTGGCTTGGTGGCGAGAG 1500
 1441 AGGCAAGCATCAAGGCAACAAACAGCTTTAGTAAATTTGGCTTGGTGGCGAGAG 1500
 1501 GAACTGTTCAATTAAGATGATTAACATTTGATACCGTAAATTTTATTTGGGCTTTC 1560
 1501 GAACTGTTCAATTAAGATGATTAACATTTGATACCGTAAATTTTATTTGGGCTTTC 1560
 1561 GTGGTGGTCTTAAATCTTAACGCGCATTCCTTAACCTTAAACGTAATCCAAATTAACG 1620
 1561 GTGGTGGTCTTAAATCTTAACGCGCATTCCTTAACCTTAAACGTAATCCAAATTAACG 1620
 1621 ACCGAGGCGCAATGATGTAACCATTAACCACTCAAGCGGTAAATGCTATTAATCTG 1680
 1621 ACCGAGGCGCAATGATGTAACCATTAACCACTCAAGCGGTAAATGCTATTAATCTG 1680
 1681 GGAACGAAGCATTTGTTCTTACTTAATGAAATTAATTAATTAATTAATTAATTAATTAAT 1740
 1681 GGAACGAAGCATTTGTTCTTACTTAATGAAATTAATTAATTAATTAATTAATTAATTAAT 1740
 1741 AAATGCTCAACAGGTTGGTGGGCAACAGATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
 1741 AAATGCTCAACAGGTTGGTGGGCAACAGATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
 1801 ACCTTATTTTAAACCAACCAACAGAGATGTAATTTTCTTCACTTCAAGTGGTCAAAAT 1860
 1801 ACCTTATTTTAAACCAACCAACAGAGATGTAATTTTCTTCACTTCAAGTGGTCAAAAT 1860
 1861 TAAAGGCGATTTACCCAAACAAAGGTAATTTTCTTCAAGGTTGAGAGCGGCAACGCG 1920
 1861 TAAAGGCGATTTACCCAAACAAAGGTAATTTTCTTCAAGGTTGAGAGCGGCAACGCG 1920
 1921 AACGCTTCAATCAATTAATTAAGTGGTCAAGATGGAATGGAAGTATTCACAGGCGAAA 1980
 1921 AACGCTTCAATCAATTAATTAAGTGGTCAAGATGGAATGGAAGTATTCACAGGCGAAA 1980
 1981 TTGTGGGATCAACATTTGATCAACCGTACATTTAAAGTGAATCTTCAATTAAG 2040
 1981 TTGTGGGATCAACATTTGATCAACCGTACATTTAAAGTGAATCTTCAATTAAG 2040
 2041 GCGGAAGTGGGTTCTTCTGCAAGTCTTCAATTTGAGGAAATGGAAGTGAAGTGAAGCA 2100
 2041 GCGGAAGTGGGTTCTTCTGCAAGTCTTCAATTTGAGGAAATGGAAGTGAAGTGAAGCA 2100
 2101 AATATGCAATGCAATTTGATGTTGTCGCAATTTTCAATTTGAGGAAATGGAAGTGAAGCA 2160
 2101 AATATGCAATGCAATTTGATGTTGTCGCAATTTTCAATTTGAGGAAATGGAAGTGAAGCA 2160
 2161 CAGATGCAAGGATTAACGATTTGCAAAAGTGAATTTAAACGATTAACGATTAACGATTAACG 2220
 2161 CAGATGCAAGGATTAACGATTTGCAAAAGTGAATTTAAACGATTAACGATTAACGATTAACG 2220
 2221 AATCTATACCAAAACCAAAATCAATGCTCTTATTAATTTAATGATGATGCAAGGCGGA 2280
 2221 AATCTATACCAAAACCAAAATCAATGCTCTTATTAATTTAATGATGATGCAAGGCGGA 2280
 2281 ATGTTAAAGGTTAGCAAACTTAATGCAATGCTCTTAAACAAATCAACGCAATTTA 2340

2281 ATGTTAAAGGTTAGCAAACTTAATGCAATGCTCTTAAACAAATCAACGCAATTTA 2340
 2341 CATTAGCAACATGCAACCAACAGGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 2400
 2341 CATTAGCAACATGCAACCAACAGGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 2400
 2401 CCGTGAATTAAGCAATTTGAACGTTAATGTCATTTAAAGATTTCAAGTCAATTTTCTT 2460
 2401 CCGTGAATTAAGCAATTTGAACGTTAATGTCATTTAAAGATTTCAAGTCAATTTTCTT 2460
 2461 TAAATAACGCAATTTTTCGCAAAATTAAGGAGCAAGGCAACAGTGAAGTGAAG 2520
 2461 TAAATAACGCAATTTTTCGCAAAATTAAGGAGCAAGGCAACAGTGAAGTGAAG 2520
 2521 AAAATGCACTGAGCAATGCTAGCATCTACATTTGAGAAATTTAAGCTTAATTAACA 2580
 2521 AAAATGCACTGAGCAATGCTAGCATCTACATTTGAGAAATTTAAGCTTAATTAACA 2580
 2581 GTACGATCAAGTTAAATTAAGCTTAATTAAGCTTAATTAAGCTTAATTAAGCTTAATTAAG 2640
 2581 GTACGATCAAGTTAAATTAAGCTTAATTAAGCTTAATTAAGCTTAATTAAGCTTAATTAAG 2640
 2641 CATTAGCAACGCAACCAACAGGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 2700
 2641 CATTAGCAACGCAACCAACAGGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 2700
 2701 AATGTAATTAAGTGGCAAGGCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAAT 2760
 2701 AATGTAATTAAGTGGCAAGGCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAAT 2760
 2761 GCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2820
 2761 GCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2820
 2821 CAGGCAAAAGCAACCGCAACCTTGAAGCAATTAATTTGTTGAAGCAAGCAATTAATTAAT 2880
 2821 CAGGCAAAAGCAACCGCAACCTTGAAGCAATTAATTTGTTGAAGCAAGCAATTAATTAAT 2880
 2881 CGTTATGATGATGATCAATTTTCTTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 2940
 2881 CGTTATGATGATGATCAATTTTCTTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 2940
 2941 GTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3000
 2941 GTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3000
 3001 AATGCAATGATTTGTAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 3060
 3001 AATGCAATGATTTGTAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 3060
 3061 TTGAACCACTGCTTAACCAACAGGTCAGCAACAGGTCAGCAACAGGTCAGCAACAGGTCAG 3120
 3061 TTGAACCACTGCTTAACCAACAGGTCAGCAACAGGTCAGCAACAGGTCAGCAACAGGTCAG 3120
 3121 GAGCAAGGTTCTGATTAACCTGCTGATTAACAGCTGTTAAAGCAATTAAGGCGCAAC 3180
 3121 GAGCAAGGTTCTGATTAACCTGCTGATTAACAGCTGTTAAAGCAATTAAGGCGCAAC 3180
 3181 AAGCTGAATGATGCTGAACCAACAAAGTAAAGGCAACAAACAAACAAACAAACAAACAAAC 3240
 3181 AAGCTGAATGATGCTGAACCAACAAAGTAAAGGCAACAAACAAACAAACAAACAAACAAAC 3240
 3241 AAGGAGCACTGTTCTGATTAACCTGCTGATTAACAGCTGTTGCAATTAAGGCGCAAC 3300
 3241 AAGGAGCACTGTTCTGATTAACCTGCTGATTAACAGCTGTTGCAATTAAGGCGCAAC 3300
 3301 TTGAGGTTATGATGCTGCAACCAATGCAAAAGTCTCTGATGCAAGGAGGAGGAGGAGG 3360
 3301 TTGAGGTTATGATGCTGCAACCAATGCAAAAGTCTCTGATGCAAGGAGGAGGAGGAGGAGG 3360
 3361 AAAAAGCAAGCAAAAGCAATTAATGCAATGCTCTTAAACAAATCAACGCAATTTA 3420

Query	Match	Similarity	Score	DB	Length
Query Match	100.0%	100.0%	DB 15;	Length 4319;	
Best Local Similarity	100.0%	Pred. No. 0;			
Matches 4319;	Conservative	0;	Mismatches	0;	Indels
					Gaps
					0;
Db	1	TCATAGTCGTTTACTAGTATTTTATATACGAAATTTACTTATTAATAATAACATTA	60		
Qy	1	TGAAAAAAGCTGTTCCTCTCTAATTTTAAACGCTGATTCATTAGGGAATAGTA	120		
Db	61				
Qy	61	TGAAAAAAGCTGTTCCTCTCTAATTTTAAACGCTGATTCATTAGGGAATAGTA	120		
Db	121	CGACAGCGTGGCTGGTCACTTTATTTTGGGATTTGATTACAAATATACCGGATTTTG	180		
Qy	121	CGACAGCGTGGCTGGTCACTTTATTTTGGGATTTGATTACAAATATACCGGATTTTG	180		
Db	121	CGACAGCGTGGCTGGTCACTTTATTTTGGGATTTGATTACAAATATACCGGATTTTG	180		
Qy	181	CCGAGATTAAGGGAATTTCAAGTGGGCTCAAAATTTAGCTTTAAACAAACAG	240		
Db	181	CCGAGATTAAGGGAATTTCAAGTGGGCTCAAAATTTAGCTTTAAACAAACAG	240		
Qy	241	GGCAATTAAGTGGCACTCATGACAAAGCCCGCATATGATTTTCTGATGTGCAC	300		
Db	241	GGCAATTAAGTGGCACTCATGACAAAGCCCGCATATGATTTTCTGATGTGCAC	300		
Qy	301	GTAACGGCGTGGCACTTGGTTGAAATCATATTTGTGAGGTGGCACTAACGTAG	360		
Db	301	GTAACGGCGTGGCACTTGGTTGAAATCATATTTGTGAGGTGGCACTAACGTAG	360		
Qy	361	GATATACAGATTTGATTTTGGTGGCAAGGGAACACCCCGATCAATCGTTTACTT	420		
Db	361	GATATACAGATTTGATTTTGGTGGCAAGGGAACACCCCGATCAATCGTTTACTT	420		
Qy	421	ATAAGATTTGAAACGAATTAATCAAAAGATATTTACATCTTTATGAGCAATT	480		
Db	421	ATAAGATTTGAAACGAATTAATCAAAAGATATTTACATCTTTATGAGCAATT	480		
Qy	481	ACCATTAATCCAGATTAATTAATTTGTTACAGACGGCTCCATTTGATATGATTCGA	540		
Db	481	ACCATTAATCCAGATTAATTAATTTGTTACAGACGGCTCCATTTGATATGATTCGA	540		
Qy	541	ATATGATGGCGTACTTATTCAGATAGAACAAATATTCGAAACGTGTTCCGATCGGCT	600		
Db	541	ATATGATGGCGTACTTATTCAGATAGAACAAATATTCGAAACGTGTTCCGATCGGCT	600		
Qy	601	CTGGAACGGCAATTTTGGCGAAATGATCAAGCAAGGCGCACAGTTGGCGGATATC	660		
Db	601	CTGGAACGGCAATTTTGGCGAAATGATCAAGCAAGGCGCACAGTTGGCGGATATC	660		
Qy	661	ATATTCGACAGTGGCAATACCAATTCAGCGCTGGAGCAGGTATGATATTTGATTT	720		
Db	661	ATATTCGACAGTGGCAATACCAATTCAGCGCTGGAGCAGGTATGATATTTGATTT	720		
Qy	721	TGGAGGCGATGTTGTTGAAACCGGAGAAATATGTCATTAACGATTCAGGCTGAAAG	780		
Db	721	TGGAGGCGATGTTGTTGAAACCGGAGAAATATGTCATTAACGATTCAGGCTGAAAG	780		
Qy	781	GGGACAGGCTGCTCCGATGTTATTTATGATGCTGAAACCAAAATGGCTTAAATG	840		

Qy	423	TATAGATTGTAAACGAATPACTACAAAAAGATAATTTACATCTTATAGAGACAT	479
Db	361	TATTAATTTGTAAACGAATPACTACAAAAAGATAATTTACATCTTATAGAGAC	420
Qy	480	TACCATATCCACGATPACTACATAATTCGTTACAGAGGGCTCAATTGATATGACTTCG	539
Db	421	TACCATATCCACGATPACTACATAATTCGTTACAGAGGGCTCAATTGATATGACTTCG	480
Qy	540	AATATGATGGGATGATCTTATTCAGATAGACAAAATTTCCAGAAAGTGTCTGATGGC	539
Db	481	CATATGATGGGATGATCTTATTCAGATAGACAAAATTTCCAGAAAGTGTCTGATGGC	540
Qy	600	TCGGAAGGAGTTTGGCAAAATGATTCAGAGCAAGGCGCAAGTTGCGGGTGCAT	659
Db	541	TCGGAAGTATGATGGGATGATTCAGAGCAAGGCGCAAGTTGCGGGTGCAT	600
Qy	660	CATTATCTGACAGCTGCAATACACAAATCAGCTGAGAGGTATATGATTCGAT	719
Db	601	AATTATTTAACAGGTGGGAATACATATATCAAAACGGTAGAGGTGATGATATTCAT	660
Qy	720	TTGGAGGGGATGTTGTAAAGGGGAAATATGTCATTCACGATTCGAGCTCAAG	779
Db	661	GTAGAGGGGATATTCGAAAGTGGGATATATGTCATTCACGATTCGAGCTCAAG	720
Qy	780	GGGAGAGTGTCTCCGATGTTATTTATGATGCTGAAAACAAAATGTTAAT	839
Db	721	GGGAGAGTGTATCTCCGATGTTATTTATGATGCTGAAAACAAAATGTTAAT	780
Qy	840	GGGATTTACGGGAGGCAACCTTTTGAGGCAAAAGAAATGAGTTCAATGGTTCG	899
Db	781	GGGATTTACGGGAGGCAACCTTTTGAGGCAAAAGAAATGAGTTCAATGGTTCG	840
Qy	900	AAATCTTATTTGATGAAATTTTGAAAGATTTTACATCATCTTACACCCGAGCT	959
Db	841	AAATCTTCTCTGATGAAATTTTACGAAAGTCAACCAATGTTTTCACCCCTAAG	900
Qy	960	GGTATGAGTGTACACAACTTAGTGAAGATATATGCTAGGGGCTATACACAGAA	1019
Db	901	GGGATGAGGCTTTATACATTTTCTAAAGTGCATGAGATGAGATGTTTCTTCCAAA	960
Qy	1020	TCAGATATCCATCAGAAATTTAATTAAGTTAGCAAAATGAGTTTACCTTTGAAAG	1079
Db	961	ATTGGAAGAAC-----TCGTGAATATCCTTTAGGCAACCAATTTAAATAGAGAT	1014
Qy	1080	AAGATTTAAGTTCAATTCATGATAGACGCACTTAATTTATCTCCAGTTTAAAC	1139
Db	1015	AAAGATCTCTATATATACGATTAATGCTCTATATATTTCTCTCAATTAAC	1074
Qy	1140	AATGGAAGAACGCTATATTTATGATCAAAACAAAGATCATTAATCTTCCGACTGAC	1199
Db	1075	AAATGGAAGAAATTTATTTTGGAGATGAAAGATTTAGATCATTAATCTTAAACGACTGAT	1134
Qy	1200	ATTACCAAGGGGCGGGTGTCTTATTTTGAAGGTAAATTTTACAGTATCTCCAAATCT	1259
Db	1135	ATGATCATAGGTGACAGCGGTTTGTATTTTGAAGGGGATTTTATATGTTTCCGCAACAA	1194
Qy	1260	AACCAATCTTGGCAAGAGCTGCAATCATGATGATGAAATAGACACCTTCTTGAAA	1319
Db	1195	AATGAAACGTGAAAGGTGCGGGCAATTCATGTCAGTGAATATGATGCGTTTGA	1254
Qy	1320	GTAAATGGGCTGGAACATGATGACTTTCTAAATTTGTTAAAGGACATTTGACGTTCAA	1379
Db	1255	GTAAACGGCTGGAATATGATGACTTTCTAAATTTGTTAAAGGACATTTGACGTTCAA	1314
Qy	1380	GCCAAAGGGGAAATTAAGTTGATCAGCGTATGAGCGATTAATGATCATTTTGGACAG	1439
Db	1315	GCCAAAGGGGAAATTAAGTTGATCAGCGTATGAGCGATTAATGATCATTTTGGACAG	1374
Qy	1440	CAGGCAAGATCAAGGCAACAAAGCTTATGTAAGTGGCTTGGTTAGGGGAGGA	1499
Db	1375	CAGGCAAGATCAAGGCAACAAAGCTTATGTAAGTGGCTTGGTTAGGGGAGGA	1434
Qy	1500	GGGAGCTGTCAATTAACGATGATTAACAAATTTGATACGCAATTAATTTATTTGGCTTT	1559
Db	1435	GGGAGCTGTCAATTAACGATGATTAACAAATTTGATACGCAATTAATTTATTTGGCTTT	1494
Qy	1560	CGTGTGTGTGCTTAATCTTAACGGGATTTCTTAACCTTTAAACGATTCGAAATAG	1619
Db	1495	CGTGTGTGTGCTTAATCTTAACGGGATTTCTTAACCTTTAAACGATTCGAAATAG	1554
Qy	1620	GNCGAGGGGCAATGATGTAACCATTAACAACTCAAGCGGATTAATGATCATTTACT	1679
Db	1555	GNCGAGGGGCAATGATGTAACCATTAACAACTCAAGCGGATTAATGATCATTTACT	1614
Qy	1680	GGGAACGAACAT---TGTTCTACTTATGCAATTAATTAATTAACCTGATTAACGA	1736
Db	1615	GGGAACGAATATATCTGCTCATCTTAATTAATTAATTAATTAACCTGATTAACGC	1674
Qy	1737	AAAGAAATTCCTTAACAGGTTGTTTGGGAAACGATTAATTAACGATGAGGCGCA	1796
Db	1675	AAAGAAATTCCTTAACAGGTTGTTTGGGAAACGATTAATTAACGATGAGGCGCA	1734
Qy	1797	TTAAACCTTATTTAAACCAACGACAGAGATGTAATTTGCTACTTTCAGGTGTACA	1856
Db	1735	TTAAACCTTATTTAAACCAACGACAGAGATGTAATTTGCTACTTTCAGGTGTACA	1794
Qy	1857	AATTTAAAGCGATATTAACCAACGATTAATTTTTCAGGCTGATACCGACA	1916
Db	1795	AATTTAAAGCGATATTAACCAACGATTAATTTTTCAGGCTGATACCGACA	1854
Qy	1917	CCGACGCTTCAATCTTTAAATTAACGTTGTGCAAAATGGAAGTATTCGCAAGGC	1976
Db	1855	CCGACGCTTCAATCTTTAAATTAACGTTGTGCAAAATGGAAGTATTCGCAAGGC	1914
Qy	1977	GAATGTGTGGATCAGATGATGATCAACGTTCAATTTAAAGCTGAATCTTCCAAAT	2036
Db	1915	GAATGTGTGGATCAGATGATGATCAACGTTCAATTTAAAGCTGAATCTTCCAAAT	1974
Qy	2037	AAAGGGAAGTGGGTTGCTTCCGATATTTTCTCAATTTGAGGCAATTTGAGACATC	2096
Db	1975	AAAGGGAAGTGGGTTGCTTCCGATATTTTCTCAATTTGAGGCAATTTGAGACATC	2034
Qy	2097	AGCATATATGCAATGCAATTTGCTGTGCAATTCACAAATATACATTTGACG	2156
Db	2035	AGCATATATGCAATGCAATTTGCTGTGCAATTCACAAATATACATTTGACG	2094
Qy	2157	CGTCAAGTTGCAAGGATTTACGACTTGTCAAAAAGTGGATTTTACCGATCAAAAGT	2216
Db	2095	CGTCAAGTTGCAAGGATTTTACGACTTGTGTAAAGGATTTTACCGATCAAAAGG	2154
Qy	2217	ATTATTTCTATACCAAAACCAAAATCAATGCTCTATATTTTATGATGATGACG	2276
Db	2155	ATTATTTCTATACCAAAACCAAAATCAATGCTCTATATTTTATGATGATGACG	2214
Qy	2277	GGCAATTTTAAAGTTTATGAAACCTTAATGCGAATGCTTATACAAATTCACAGCA	2336
Db	2215	GGCAATTTTAAAGTTTATGAAACCTTATGCTTAATGCTTATATATATATACGCA	2274
Qy	2337	TTTACATTAACCAATGCGACCAAAATGAGCAATTTGCACTTTCCGCAATTTCACT	2396
Db	2275	TTTACATTAACCAATGCGACCAAAATGAGCAATTTGCACTTTCCGCAATTTCACT	2334
Qy	2397	-----	2396
Db	2335	GCAAGGTTAATTAATGCACTTAAATGGGCGATGATGATTTAGCGGATATACGCTTTT	2394
Qy	2397	-----GCA	2399
Db	2395	ACATTAAGCAATGAGCAACAGATTTGCAATCATGCTTCAATCAGCAAGCTCAAGCA	2454
Qy	2400	ACGATTAATTAATGCAATTTGAACGCTATGCTTATTAACGATTTCACTCAATTTCT	2459
Db	2455	ACGATTAATTAATGCAATTTGAACGCTATGCTTATTAACGATTTCACTCAATTTCT	2514
Qy	2460	TTAAAAACAGCATTTTTGGACCAAAATTCAGGAGACAAAGGCAACAGATGATGCTTT	2519

Db 2515 TTTAAAAACATGCTATTTTTCACCAATTCAGGCGCAAGAAACACACAGTGCCTG 2574
 Qy 2520 GAAATGCGACTTGGACATAGCTTACCGATCTACATTCGAGATTTAAGCTAAATAC 2579
 Db 2575 GAAATGCGACTTGGACATAGCTTACCGATCTACATTCGAGATTTAAGCTAAATAC 2634
 Qy 2580 AGTACATCAGCTTAAATTCAGCTTATTCAGCTTACCTCAACATTCAGCTTAAAT 2636
 Db 2635 AGTACATCAGCTTAAATTCAGCTTATTCAGCTTACCTCAACATTCAGCTTAAAT 2694
 Qy 2637 CGTTTATGAGAGGAAACACAGCAATTCAGCTTACCTCAACATTCAGCTTAAAT 2696
 Db 2695 CGTTTATGAGAGGAAACACAGCAATTCAGCTTACCTCAACATTCAGCTTAAAT 2754
 Qy 2697 GTAAATGTAATGAGGCGACAGGCAATTCATTCATTCATTCATTCATTCATTCAT 2756
 Db 2755 GTAAATGTAATGAGGCGACAGGCAATTCATTCATTCATTCATTCATTCATTCAT 2814
 Qy 2757 AAAAGCGATTAATTTAAATTTTCCATTCAGCTTACCTCAACATTCAGCTTAAAT 2816
 Db 2815 GAAAGCGATTAATTTAAATTTTCCATTCAGCTTACCTCAACATTCAGCTTAAAT 2874
 Qy 2817 AACACAGCAAGAACCCGAAACCTTACAGCAATTCATTCATTCATTCATTCATTCAT 2876
 Db 2875 AACACAGCAAGAACCCGAAACCTTACAGCAATTCATTCATTCATTCATTCATTCAT 2934
 Qy 2877 CAACCGTTACAGTATGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 2936
 Db 2935 CAACCGTTACAGTATGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 2994
 Qy 2937 TTACGTTAAATTTAGTGAAGATTCAGCTTACCTCAACATTCAGCTTAAAT 2996
 Db 2995 TTACGTTAAATTTAGTGAAGATTCAGCTTACCTCAACATTCAGCTTAAAT 3054
 Qy 2997 CAGGAATGCAATGATTTAGTGAAGATTCAGCTTACCTCAACATTCAGCTTAAAT 3056
 Db 3055 CAGGAATGCAATGATTTAGTGAAGATTCAGCTTACCTCAACATTCAGCTTAAAT 3114
 Qy 3057 CAAGTTGAACGATGCTTAAACCAACAGGTGAGCCAAAGTGGCTCAAGAGCA 3116
 Db 3115 CAAGTTGAACGATGCTTAAACCAACAGGTGAGCCAAAGTGGCTCAAGAGCA 3168
 Qy 3117 GCGAGAGAGCGCTTCTGATACCTGCTGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 3176
 Db 3169 --AGACGGGTGTCTGATACCTGCTGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 3225
 Qy 3177 AAACAGCTGAATGCTGCTGAACCAACAGGTGAGCCAAAGTGGCTCAAGAGCA 3236
 Db 3226 AAACAGCTGAATGCTGCTGAACCAACAGGTGAGCCAAAGTGGCTCAAGAGCA 3285
 Qy 3237 TCAAAAAGAC-----AGTGTTCCTGATCCCTGCTTCAAGCTTAAAGCTTAAAGCT 3284
 Db 3286 TCAAAAAGAC-----AGTGTTCCTGATCCCTGCTTCAAGCTTAAAGCTTAAAGCT 3345
 Qy 3288 GCATTAAAGCGGCACTTGAATGATTCAGCTTACCTCAACATTCAGCTTAAAT 3344
 Db 3346 GCATTAAAGCGGCACTTGAATGATTCAGCTTACCTCAACATTCAGCTTAAAT 3405
 Qy 3348 GCTCAAGAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 3401
 Db 3406 GCTCAAGAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 3465
 Qy 3408 AATAGTGCCTTATCAGAAATTCCTGCAACAGTAAATAGTATCTTCTGTTAAGATGA 3461
 Db 3466 AATAGTGCCTTATCAGAAATTCCTGCAACAGTAAATAGTATCTTCTGTTAAGATGA 3525
 Qy 3468 TTATGCTGCTTTTGTAGATCAGCAATTCAGCTTACCTCAACATTCAGCTTAAAT 3521
 Db 3526 TTATGCTGCTTTTGTAGATCAGCAATTCAGCTTACCTCAACATTCAGCTTAAAT 3585
 Qy 3528 AAAAGCGATTAATTTAAATTTTCCATTCAGCTTACCTCAACATTCAGCTTAAAT 3581

Db 3586 AAAAGCGATTAATTTAAATTTTCCATTCAGCTTACCTCAACATTCAGCTTAAAT 3642
 Qy 3588 CAAATTTGGGATGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 3641
 Db 3643 CAAATTTGGGATGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 3702
 Qy 3645 CTTCAAGATTAATTCCTTGAAGCAAGGTTAAATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 3701
 Db 3703 CTTCAAGATTAATTCCTTGAAGCAAGGTTAAATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 3762
 Qy 3705 GCTTATGAGAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 3761
 Db 3763 GCTTATGAGAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 3822
 Qy 3765 AGTGCAGTAAATTCCTTGAAGCAAGGTTAAATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 3821
 Db 3823 AGTGCAGTAAATTCCTTGAAGCAAGGTTAAATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 3882
 Qy 3825 GCGTGAATGCAAGTATTCAGCTTACCTCAACATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 3881
 Db 3883 GCGTGAATGCAAGTATTCAGCTTACCTCAACATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 3942
 Qy 3885 GTAAATGCTATTTATTCAGCTTACCTCAACATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 3941
 Db 3943 GTAAATGCTATTTATTCAGCTTACCTCAACATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 4002
 Qy 3945 CTTACGCTGATTTATTCAGCTTACCTCAACATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 4001
 Db 4003 CTTACGCTGATTTATTCAGCTTACCTCAACATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 4062
 Qy 4005 ACAGATTAATTCAGCTTACCTCAACATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 4061
 Db 4063 ACAGATTAATTCAGCTTACCTCAACATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 4122
 Qy 4065 AACGTAACCAACGTTAAATTCAGCTTACCTCAACATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 4121
 Db 4123 AACGTAACCAACGTTAAATTCAGCTTACCTCAACATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 4182
 Qy 4125 AAAGAGTGGATTTAAAGGCAAAATTCAGCTTACCTCAACATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 4181
 Db 4183 AAAGAGTGGATTTAAAGGCAAAATTCAGCTTACCTCAACATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 4242
 Qy 4185 TCTCAAGGTTCAACATTCAGCTTACCTCAACATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 4241
 Db 4243 TCTCAAGGTTCAACATTCAGCTTACCTCAACATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAATTCAGCTTAAAT 4302
 Qy 4245 TCA 4244
 Db 4303 TCA 4305

RESULT 5
 US-10-329-960-1/c
 ; Sequence 1, Application US/10329960
 ; Publication No. US2003009277A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fleischmann et al.
 ; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
 ; TITLE OF INVENTION: Thereof, and Uses Thereof
 ; FILE REFERENCE: FBI86P1
 ; CURRENT APPLICATION NUMBER: US/10/329,960
 ; PRIOR FILING DATE: 2003-01-02
 ; PRIOR APPLICATION NUMBER: US 09/643,990
 ; PRIOR FILING DATE: 2000-08-23
 ; PRIOR APPLICATION NUMBER: US 08/487,429
 ; PRIOR FILING DATE: 1995-06-07
 ; PRIOR APPLICATION NUMBER: US 08/426,787
 ; NUMBER OF SEQ ID NOS: 1
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 1630121
 ; TYPE: DNA

ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40868)..(40868)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102596)..(102596)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

Query Match	69.9%;	Score 3019.8;	DB 15;	Length 1830121;
Best Local Similarity	83.0%;	Pred. No. 0;		
Matches 3641:	Conservative	2;	Mismatches 650;	Indels 93; Gaps 14

Db	2781.37	GATTAATCATAGAGCTCGTTTACATTAATTGGAACTGAAAGCTGAAAGCTGTGGGATATACG	2783048
Oy	531	TCGAATTAATTAATGCGACTCATTTACAGATAGAAACAATATACAGAGGTGGTGTATC	596
Db	27804.7	ACAAATATAGATGAAAGAAAGTATATGTCTGATAGAGAGAACTATCTGAGGTGTACGTATA	277898
Oy	597	GGCTGTGACGGCAGTTTGGGGAAATGATCAAGACAAGCGC-----	639
Db	277987	GGCTCAGAGACGTACAGTATGGGGTACAGATTAAGATGAAGAAAGAAATGATCATAGTTCA	277928
Oy	640	-ACCAAGTTGCCGTGCTATTCATTTATGTGACAGCTGGCAATTAACACAAATACAGGTGGA	658
Db	277927	TATTAATGTCTGAGGTGCATATGATATCTTACTCGACGAAATATCCATCTACTCAGGTGGA	277868
Oy	699	GCAGGTATATGATTCGTATTTGGGAGGGAGTGTCCGTAAGAGGGAGATATGTTCTCA	758
Db	277867	AATGTATATGTAACAGTCATCTGTAGGTGTAATGTATAGTTCAGCTTATCTATGTGTCA	277808
Oy	759	TTACCATGATGAGGCTCAAAAGGGGACAGTGTCTTCGATGTTATTTATGATGCTGAA	818
Db	277807	TTACCAAGCGGTGTTCTTAAGGGCATAGCGGTTCCGCAAGTTTATTTATGATGCTGAG	277748
Oy	819	AAACAAAATGGTTAATATATGAGTATATAGGGAATATAGGAGCAACCTTTTGAGGCAAGAA	878
Db	277747	AAGAAACATGGCTTATTAATATGCTGTATTAACAACTGGGCATCTTTTTCGAAAGCGT	277688
Oy	879	AATGGGTTCAATGTGTTCCAAATCTTATTTTGAAT-----GAAATTTTC	923
Db	277687	AATGGGTTCAAGTTAATATAGTGAAGATGGTTTATATATGAAAGTCTTGCGGTTGATACC	277628
Oy	924	GAAAGAGATTTACATCATCTTACACCCGACCTGTGTAATGAGAGTATGACAAATTAGT	963
Db	277627	CTTAGTGTGTTTTCACAGCTATATATCCCAATTAATGACATATTTCTTGTATCAAT	277568
Oy	984	GAAATGATATATGTCAGGGGTCATATATCTACACAAATCAGAAATACATCAGAAATTAA	1043
Db	277567	AATGATGTACGGTAATTAATTAACTTAACTAGACCTAGTAAGATGGCTTTAAACAAA	277508
Oy	1044	ATTACGTTAGCAATATGAGTTTACCT-----TGAAAGAGAGATTAAGTTTCAAT	1094
Db	277507	TCGAAGTGTGAACTGTGAAGTATTTATTCATCTGTTAAATCAACCGGTAAAGACAT	277448
Oy	1095	AATCTTAGATAGAGAGACCTAAATTTATTTCCACAGTTTAAACATGAGAGAAAGCTA	1154
Db	277447	CTTAAGACAGACGAGGCTATATATTTTACCAACAGAAATGATATGAGAAAAATTT	277388
Oy	1155	TATTTATGATCAAAACAGAGATCATTTCTTCGATCTGACATTAACCAAGGAGCG	1214
Db	277387	TACCTAGGAGCAAGAAAGAAAGAACTTTAACATGAAATATATTAATCAAGTGTCT	277328
Oy	1215	GGTGTCTTTATTTTGAGGGTAATTTTACAGATCTCCAA--ATTGTACCAAACTGG	1271
Db	277327	GGTGAATTTACCTTTGAAGGTAAATTTTGTGTAAAGCAGCAAAATATATTAACCTGG	277268
Oy	1272	CAAGAGCTGAGCATCATCATATGAGAAATAGACCGTACTCTGAAAGTAATGSCGTG	1331
Db	277267	CAGGTGTGAGGGTATCTATGAGACAAATGCAACCTGTTGATATGAAAGTCAACATCTCT	277208
Oy	1333	GAACTATATGACTCTTTCAAATATGGTTAAGAGAACTATGCAAGCTTCAAGCCAAAGGGAA	1391
Db	277237	GAAATATATCTTTATCTAAATATGGTATAGGCACTTTATTTAGTCATGTGTAAAGGAAG	277148
Oy	1392	AATTAAGTTGCATAGAGGTAGGCAGATGTAAATCATTTTGTAGACAGAGGAGACAT	1451
Db	277147	AATTTAGGAATTTAATGTGGGTAAAGGCAAAATCATTTCTAGATCAACAGAGATGAA	277088
Oy	1452	CAAGCACAACACAGCCTTTACTGAAATGGCTTGGTTAGCGGACAGAGGACCTGTTCA	1511
Db	277087	GGCGGTCAAAACAGCTTTCAAGAAAGTGGCACTGTAGCGGTGACAGCAAGTTTCA	277028
Oy	1512	TTAATGATGATTAACATTTGATACCAATAAATTTATTTCCGCTTCTGCTGTGGTGC	1571
Db	277027	TTAATATGTAAGATCAAGTATGTTCTTACATATCTATTTCCGATTTGCTGTGGTGC	276968


```
Db 274828 TATCAATGGGGGCAATTTCATTTGGTGTAAATGCGTGGGGAATAGTSCGAGTAAA 274769
Qy 3774 ATGGCTGAAGAACAAAGCCGAAATTCATTCGAAAGGAAATTAATGCGTGAATGCA 3833
Db 274768 ATGGCTGAAGAACAAAGCCGAAATTCATTCGAAAGGAAATTAATGCGTGAATGCA 274709
Qy 3834 AGTTATCAGTTCGCTTAGGGCAATTTGGGCAATTCAGCTTATTTTGGAGTTAATGCTAT 3893
Db 274708 AGTTATCAGTTCGCTTAGGGCAATTTGGGCAATTCAGCTTATTTTGGAGTTAATGCTAT 274649
Qy 3894 TTTATTCAGCGTGAATTAATTCATTCGAGAGAGAGAGTGAAGAGCTTACGCTTGCA 3953
Db 274648 TTTATTCAGCGTGAATTAATTCATTCGAGAGAGAGAGTGAAGAGCTTACGCTTGCA 274589
Qy 3954 TTTATTCAGCGTGAATTCAGCTTATTCATTCGAGAGAGAGTGAAGAGCTTATTC 4013
Db 274588 TTTATTCAGCGTGAATTCAGCTTATTCATTCGAGAGAGAGTGAAGAGCTTATTC 274529
Qy 4014 AGCGTTAAGCTTATTTCTTCGTCATTTATGTTGATGTTTCAAGCGTACGTAACCAACC 4073
Db 274528 AGCGTTAAGCTTATTTCTTCGTCATTTATGTTGATGTTTCAAGCGTACGTAACCAACC 274469
Qy 4074 ACCGTTAATCTCAGCGTGTGCAACCACTTTGCACTTATTTGGCAAAAAGAGTGGGA 4133
Db 274468 ACTGTAATAGCAAGCATGTTCGCAACAATTTTGGGCGTATTTGGCAAAAAGAGTGGGA 274409
Qy 4134 TTAAGGCAAGAAATTTTACATTCGAAATTTTCGCTTATTCGCAAAATCTGAAGTGTCA 4193
Db 274408 TTAAGGCAAGAAATTTTACATTCGAAATTTTCGCTTATTCGCAAAATCTGAAGTGTTCG 274349
Qy 4194 CAATCGGCAAAACAGCAAAATGTGGGCGTGAATTTGGGCTATCGTTGTTAAATCAACA 4253
Db 274348 CAATCGGCAAAACAGCAAAATGTGGGCGTGAATTTGGGCTATCGTTGTTAAATCAACA 274289
Qy 4254 TAAATTTATCGTTTATTTATTAACAGAGTGGTCAATCAATCCCACTTTTATTC 4313
Db 274288 TAAATTTATCGTTTATTTATTAACAGAGTGGG-----GCAAGTCCCACTTTTATTC 274234
Qy 4314 AATAAT 4319
Db 274233 AATAAT 274228

RESULT 6
US-10-329-670-1/c
; Sequence 1, Application US/10329670
; Publication No. US20040018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; FILE REFERENCE: PH186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
```

```
/ LOCATION: (9921)..(9921)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (10150)..(10150)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (29298)..(29298)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (36543)..(36543)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (36551)..(36551)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (36536)..(36536)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (40838)..(40838)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (44416)..(44416)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (44905)..(44905)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (44975)..(44975)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (45593)..(45593)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (45732)..(45732)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (47036)..(47036)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (51334)..(51334)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (51602)..(51602)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (51786)..(51786)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (51805)..(51805)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (55369)..(55369)
/ OTHER INFORMATION: n equals a, t, g or c
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (65309)..(65309)
```

```

1  NAME/KEY: misc_feature
2  LOCATION: (145058)..(145058)
3  OTHER INFORMATION: n equals a, t, g or c
4  FEATURE:
5  NAME/KEY: misc_feature
6  LOCATION: (145171)..(145171)
7  OTHER INFORMATION: n equals a, t, g or c
8  FEATURE:
9  NAME/KEY: misc_feature
10 LOCATION: (145942)..(145942)
11 OTHER INFORMATION: n equals a, t, g or c
12 FEATURE:
13 NAME/KEY: misc_feature
14 LOCATION: (147197)..(147197)
15 OTHER INFORMATION: n equals a, t, g or c
16 FEATURE:
17 NAME/KEY: misc_feature
18 LOCATION: (150841)..(150841)
19 OTHER INFORMATION: n equals a, t, g or c
20 FEATURE:
21 NAME/KEY: misc_feature
22 LOCATION: (152500)..(152500)
23 OTHER INFORMATION: n equals a, t, g or c
24 FEATURE:
25 NAME/KEY: misc_feature
26 LOCATION: (152530)..(152530)

```

Query Match	69.9%	Score 3019.8;	DB 16;	Length 1830121;
Best Local Similarity	83.0%;	Pred. No. 0;		
Matches 3641; Conservative	2;	Mismatches 650;	Indels 93;	Gaps 14

Qy	1	CAATGTCGCTACGA-GATATTTTAAAGGAAATATCTAATTAATAAGCT	59
Ds	278567	TCATATGTCGTTTACCAAGCATTTTTTAAATGCAAAATATCTAATTAATAAGCAT	278528
Qy	60	ATGAAAAAACTGATTTTGGCTTAATTTTAAACCGCTTGATTCATTAGGATAGTA	119
Ds	278527	ATGAAAAAACTGATTTTGGCTTAATTTTAAACCGCTTGATTCATTAGGATAGTA	278468
Qy	120	TCGCAAGCGTGGGCTGCTCAACATTAATTTGGATGATTACCAATATTTTCGTATTT	179
Ds	278467	TCGCAAGCGCTGGGAGGTCTACTTAATTTGGATGATTACCAATATTTTCGTATTT	278408
Qy	180	GCCGACATAAAGGAAAGTTGACAGCTTGGGGCTCAAAATATTAAGGTTTAAACAAACA	239
Ds	278407	GCCGACATAAAGGAAAGTTGACAGCTTGGGGCTCAAAATATTAAGGTTTAAACAAACA	278346
Qy	240	GGGCAATTAAGTTGGCACATCAATGACAAAGCCCGATGATTTTCTGATGTCA	299
Ds	278347	GGGCAATTAAGTTGGCACATCAATGACAAAGCCCGATGATTTTCTGATGTGTG	278286
Qy	300	CGTAAAGCGGTGCGACGCTTGTTGTAATATTAATTTGGAGCGTGGACATTAACGA	359
Ds	278287	CGTAAAGCGGTGCGAGCATTAAGTAGCATATGATTAATTTGGAGCGTGGACATTAACGAC	278228
Qy	360	GGATATACAGATGTTGATTTGGTGCAGAGGGAAACACCCCGATCAACATCCGTTTACT	419
Ds	278227	GGATATACAGATGTTGATTTGGTGCAGAGGGAAACCCGATCAACACCCGCTTACT	278168
Qy	420	TATPAGATTTGAAACGAAATTAATCTACAA--AAAGATTAATTACATCTTATTAAGAC	476
Ds	278167	TATCAAAATGTTAAAGAAATTAATTAATCACTTGGGAGAAAGACATCCCTTAATGATGA	278108
Qy	477	GATTACCAATACACAGATTACGTAATAATGCTTAACGAAGGGGCTCCAAATTAATGATGCT	536
Ds	278107	GATTATACATATGCTCGTTGCTATTAATTTGTAACTGAAGCTGAACCTGTGGGTATGACA	278048
Qy	537	TCGAAATGATATGCAAGTACTTAATTCAGATGAAACAAATATCCAGAACGTGTTGCTATC	596
Ds	278047	ACAAATATGATGCAAAAGATATATGCTATGAGAGACATACCTGAGCGCTGTAGCTATA	277988
Qy	597	GGCTCTGACGCGAGCTTTTGGGAAATATCATCAAGACAAAGGCG-----	639

Db 277987 GAGCTGACGCTGAGTATGGCGTACAGATTAAGATGAAGAAAGCATGTACATAGTTCA 277928
 Qy 640 -ACCAAGTCCCGGTGATATCATATATCTGACAGCTGCGAATACACATCAGCGTGA 698
 Db 277927 TATATGCTCAGGTGCTATATGCTTATCTTACGAGGAAATACCACTAGCGTGA 277868
 Qy 699 GCGAGTATGATATTCGTATTTGGAGGCGATGTTCTGTAAGCGGAGAAATGATGCA 758
 Db 277867 AATGATATGCTACATCAATCTAGTGTAAAGTATGACCTTATCATATATGTTCA 277808
 Qy 759 TTACGANTGAGGCTCAAGGCGGACAGTGTCTCCGATGTTATTTATGATCTGA 818
 Db 277807 TTACCAAGGCGTCTTAAAGCGATAGCGGTGCGCAATGTTATATATGATCGAAG 277748
 Qy 819 AAACAAAATGTTATATTAATGGAATCTACGAGAAAGCAACCTTTGAGGCAAGAA 878
 Db 277747 AAGAAACATGCTTATATATGCTATTTACAACTGCGGCACTCTTTTTCGAAAGAGT 277688
 Qy 879 AATGGTTCAATGTTGCGAAATCTTATTTGAT-----GAAATTTC 923
 Db 277687 AATGGTTCAATGTTGCGAAATCTTATTTGAT-----GAAATTTC 923
 Qy 924 GAAAGATTTACATACATCACTTACACCCGAGCGGTATGAGATGACAAATTAGT 983
 Db 277627 CTTATGTTTTCAGCTATATTTCCCAATTAATGACATTTTCCTTGTATCAAT 277568
 Qy 984 GGAATGATATGCTCAGGCGGTCTATATCTCAGAAATCAGAAATACATCAGAAATTA 1043
 Db 277567 AATGATGTTACGTTAAATTTAACTTTACAGCACTAGTAAAGTGGCTTAAAGCAAA 277508
 Qy 1044 ATTAAGTTAGCAATTAATGATTTACT-----TTGAAGAGAGATTAAGTTAT 1094
 Db 277507 TCAGATGTAAGATGTAAGTATTTATTAATCACTGTTAATGAACCGCTTAAGACAT 277448
 Qy 1095 AATCTGATTAAGAGGACCTAATTTATTTCTCCAGTTTAAACAATGAGAAAGCTA 1154
 Db 277447 GTTAAAGCAGCAGCGCTATATATTTTACCAACCAATGAGATTTGAAAAATAT 277388
 Qy 1155 TATTTATGATCAAAAACAAGATCATTAATCTGATGATGCACTTAACCAAGGCGG 1214
 Db 277387 TACTTGGGACCAAGGAAAGAAAGAACTTTACATGCAAAATATTAATGATAGGTC 277328
 Qy 1215 GGTGGTTTATTTTGAAGGTATTTTACATGATCTCA--ATTCAACCAAACTTG 1271
 Db 277327 GGTGGATTAATCTTGAAGGTATTTTGTGTAAGGCAAGCAAAATATTAATCTTG 277268
 Qy 1272 CAAGAGCTGGGATACATGTAAGTAAATACACCGTACTTGAAGTAAATAGCGCTG 1332
 Db 277267 CAAGGTGAGGCGTATGTAAGCAAGATGCACTGTATATGAAAGTTCAATCT 277208
 Qy 1332 GAACATGATCGATTTCTTAAATTTGTAAGGACATGCACTTCAACCAAGGCGAA 1391
 Db 277207 GAATATGATGTTATCTTAAATTTGTAAGGCACTTATATGATCAATGTAAGGAAAG 277148
 Qy 1392 AATTAAGTTGATGAGGTGAGGATGAGTAAAGTCAATTTTGAAGCAAGGCGACAT 1451
 Db 277147 AATTTAGAGGTTTAAAGCGGGAACGCAAGATCTTCAATCAACAGCAGATGA 277088
 Qy 1452 CAAGCAACAAACAGCTTTAGTAAATTTGCTGTGAGGCGAGAGGACCTGTCA 1511
 Db 277087 GCGGCTCAAAACAGCTTTCAAAAGGTTGAGGCGGTGAGGCAACGTTCA 277028
 Qy 1512 TTAAAGATGATTAACATTTGATGCAATTTTATTTGCTTTGCTGTGCTG 1571
 Db 277027 TTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 276968
 Qy 1572 TTGATGCTTAAGGCGATGATTAACCTTAAAGGTAATCAAAATCGAGAGGCGCA 1631
 Db 276967 TTGATGCTTAAGGCGATGATTAACCTTAAAGGTAATCAAAATCGAGAGGCGCG 276908
 Qy 1632 ATGATTTGATCAATTAACACTCAAGCGGCTATGCTATCTAGGAAAGCAAGC 1691
 Db 276907 ATGATTTGATCAATTAACACTCAAGCGGCTATGCTATCTAGGAAAGCAAGC 1691
 Qy 1692 AT---TGTTACCTCAATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1748
 Db 276847 ATTAATGCTTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276788
 Qy 1749 TACACGTTGTTTGGCGAAACAGTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1808
 Db 276787 TACACGTTGTTTGGCGAAACAGTAAATTAATTAATTAATTAATTAATTAATTAAT 276728
 Qy 1809 TATTAACCAACCAAGAGTGTACTTCTGCTACTTCAAGGTGATCAATTAATTAATTAAT 1868
 Db 276727 TATTAACCAACCAAGAGTGTACTTCTGCTACTTCAAGGTGATCAATTAATTAAT 276668
 Qy 1869 GATATTAACCAACCAAGAGTGTACTTCTGCTACTTCAAGGTGATCAATTAATTAATTAAT 1928
 Db 276667 GATATTAACCAACCAAGAGTGTACTTCTGCTACTTCAAGGTGATCAATTAATTAAT 276608
 Qy 1929 AATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1988
 Db 276607 AATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276548
 Qy 1989 GATCAAGATGATCAACCGTATTTAACTGTAATTAATTAATTAATTAATTAATTAAT 2048
 Db 276547 GATTAAGATGATCAACCGTATTTAACTGTAATTAATTAATTAATTAATTAATTAAT 276488
 Qy 2049 GCGGTGTTTTCGCAATGTTCTCAATGAGGAAATTTGACAGTCAAGCAATTAATGA 2108
 Db 276487 GCGGTGTTTTCGCAATGTTCTCAATGAGGAAATTTGACAGTCAAGCAATTAATGA 276428
 Qy 2109 AATGCCAATTTGTTGTTGCGAAATCAACAAATTAATTAATTAATTAATTAATTAAT 2168
 Db 276427 AATGCCAATTTGTTGTTGCGAAATTAACAAATTAATTAATTAATTAATTAATTAAT 276368
 Qy 2169 ACGATTAACGATTTGTAAGATGATTTAACGATTAATTAATTAATTAATTAATTAAT 2228
 Db 276367 ACGATTAACGATTTGTAAGATGATTTAACGATTAATTAATTAATTAATTAATTAAT 276308
 Qy 2229 CCAAAACCAATCATGCTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2288
 Db 276307 CCAATTAACCAATCATGCTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276248
 Qy 2289 GGTTAAGAACTTAATGCAATGCTCACTTAATTAATTAATTAATTAATTAATTAATTAAT 2348
 Db 276247 GGTTAAGAACTTAATGCAATGCTCACTTAATTAATTAATTAATTAATTAATTAAT 276188
 Qy 2349 AATATGCAACCAACCAATGCAATGCTCACTTAATTAATTAATTAATTAATTAATTAAT 2408
 Db 276187 AATATGCAACCAACCAATGCAATGCTCACTTAATTAATTAATTAATTAATTAATTAAT 276128
 Qy 2409 AATGCAACCTGAACGATTAATGATTTAACGATTAATTAATTAATTAATTAATTAAT 2468
 Db 276127 AATGCAACCTGAACGATTAATGATTTAACGATTAATTAATTAATTAATTAATTAAT 276068
 Qy 2469 AGCATTTTTCGCAATTAATGAGGAGCAAAAGGCAACAGTACGTTGAAAAATGCG 2528
 Db 276067 AGCATTTTTCGCAATTAATGAGGAGCAAAAGGCAACAGTACGTTGAAAAATGCG 276008
 Qy 2529 ACTTGAAGATGCTGAGATGCTACATGAGATTAATTAATTAATTAATTAATTAATTAAT 2588
 Db 276007 ACTTGAAGATGCTGAGATGCTACATGAGATTAATTAATTAATTAATTAATTAATTAAT 275948
 Qy 2589 AGCTTAATTTGCTATTAATGATGCTCAACATTAATGCTGCT--CGCGCTTATTA 2645
 Db 275947 AGCTTAATTTGCTATTAATGATGCTCAACATTAATGCTGCT--CGCGCTTATTA 275888
 Qy 2646 GAGAGGAAACCAAGGCAATCGGAGAGATGCTTCAACATTAATTAATTAATTAATTAAT 2705
 Db 275887 GAGAGGAAACCAAGGCAATCGGAGAGATGCTTCAACATTAATTAATTAATTAATTAAT 275828
 Qy 2706 AATTAAGTGGGCAAGGCAATCGGAGAGATGCTTCAACATTAATTAATTAATTAATTAAT 2765
 Db 275827 AATTAAGTGGGCAAGGCAATCGGAGAGATGCTTCAACATTAATTAATTAATTAATTAAT 275768

QY	2766	AAATTAATAATATCCAAATGACGCTGAGGGCCGATTACATATATCTGTCCAAACAAGCC	2825
Db	275767	AAATTAATAATATCCAAATGACGCTGAGGGCCGATTACATATCTGTCCAAACAAGCC	275708
QY	2826	AAAGAACCCGAAACCTTGAGCAATTACCTTGGTTGAAGGAAAGATATCAACCGTTA	2885
Db	275707	AAAGAACCTTGAGCAACCTTGAGCAATTACCTTGGTTGAAGGAAAGATATCAACCGTTA	275648
QY	2886	TCAGATAGCGTCAATTTACTCTTATGAAAATGACACGTTGATGACAGTGCATTAAGTTAT	2945
Db	275647	TCAGATAGCGTCAATTTACTCTTATGAAAATGACACGTTGATGACAGTGCATTAAGTTAT	275588
QY	2946	AAATTAAGTAAGAATGATGCGAATTCGCGTTGCAATAACCAATTTAAAGACAGCAATTG	3005
Db	275587	AAATTAAGTAAGAATGATGCGAATTCGCGTTGCAATAACCAATTTAAAGACAGCAATTG	275528
QY	3006	CACATGATTTTGTATAGAGAGAGAGAGAGAGACAACTATGAAGCCAAACAGTTGA	3065
Db	275527	CTCAATGATTTTGTATAGAGAGAGAGAGAGAGACAACTATGAAGCCAAACAGTTGA	275468
QY	3066	CCGACGTCTTAAACACAAACAGGTGAGCCAAATGCGGTCAAGAGAGAGAGAGAGA	3125
Db	275467	CAGCTCTCTGAAAAACAAAAAGTAAAGCCAAAGCGCGTCAAG-----MAGCG	275417
QY	3126	GCCTTCTGATACCTGCTGATCCAAAGCCTGTTAAAGCATTAAGAGCCAAACAGCT	3185
Db	275416	GTGTGTCTGATACCCGCTGCTCAAGACCTGTTAAAGCATTAAGAGCCAAACAGCTT	275357
QY	3266	GAACTACGCTGAAAAACAAAAAGTAAAGCAAAAAACAAAAAGTGGCGGTCAAAAAA	3245
Db	275356	GAACTACGCTGAAAAACAAAAAGTAAAGCAAAAAACAAAAAGTGGCGGTCAAAAAA	275297
QY	3246	-----GCAGTGTCTTCTGATCCCTGCTTATCAAGAGCTGTT---CGCATTAAGA	3293
Db	275296	GCATGAGTGCAGCGTCTTCTGATACCCGTTGATCTTAAGCAGTTAAAGATTCGA	275237
QY	3294	GGCGACCTTGAGGTTATTTATGCCCCAGCAATCGAAAAAGATCGTCTAGCTCAAGA	3353
Db	275236	GTCMAACTTGAAGGTTATTAATGCCCCAACCGCAATGAAAAAGAACTCAAGATCAAGG	275177
QY	3354	GAAGCGAAAAACAAACGCAACAAAAAGACTGATCAGCGCTTATCAAAATAGCTGGTTA	3413
Db	275176	G-----AACAGGCAACAAAAAGATGATCAGCGCTTATCAAAATAGCTGGTTA	275126
QY	3414	TCGAAATATCTGCAACAGTAAATAGTATGCTTCTGTTCAAGATGAATTAGTCGTCTT	3473
Db	275125	TCGAGATGTCGCAACAGTAAATAGTATGCTTCTGTTCAAGATGAATTAGTCGTCTT	275066
QY	3474	TTTGTAGATCAAGCAACATTCGCGCTGTGACAAATATGCGACAGATTAAGACGCTAT	3533
Db	275065	TTTGTAGATCAAGCAACATTCGCGCTGTGACAAATATGCGACAGATTAAGACGCTAT	275006
QY	3534	GATTCGTGATGCGTTCCGCTTATCAGCAGCAGAAAAACAACTTACGCTAAATTGGAGT	3593
Db	275005	GATTCGTGATGCGTTCCGCTTATCAGCAGCAGAAAAACAACTTACGCTAAATTGGAGT	274949
QY	3594	CAAAAACCTTATGCTAATGACGAATTTGGAGCAGTTTTCTCGCATAGCGCTTCAGATAT	3653
Db	274948	CAAAAACCTTATGCTAATGACGAATTTGGAGCAGTTTTCTCGCATAGCGCTTCAGATAT	274889
QY	3654	ACCTTTGATGAACAGGTTAAAAATCAACGCAATTAACAGATAGTGGGGTTTGGCCAA	3713
Db	274888	ACCTTTGATGAACAGGTTAAAAATCAACGCAATTAACAGATAGTGGGGTTTGGCCAA	274829
QY	3714	TATCAATGGAGCGAATTTACAAATTTGGTGTAAACGTGGGAACGGGAATCAGTGCAGTAA	3773
Db	274828	TATCAATGGAGCGAATTTACAAATTTGGTGTAAATGTGGGTGCGGGAATTAAGTGCAGTAA	274769
QY	3774	ATGGCTGAAGAAACAAGCGCAAAATATCTATGAAAAGCATTAATATATGCGCTGAATCA	3833
Db	274768	ATGGCTGAAGAAACAAGCGCAAAATATCTATGAAAAGCATTAATATATGAGAGATCA	274709

QY	3884	AGTATCACTGCTCGTTATGGGCAATTGGGCACTCACCTTAATTGGAGTATAGCAT	3893
Db	274708	AGTATCACTGCTCGTTATGGGCAATTGGGCACTCACCTTAATTGGAGTATAGCAT	2746498
QY	3894	TTTATGGAAGTGAATAATTATCATCTGAGGAAGTAGAGTGAACCCCTAGCCTTGA	3953
Db	274648	TTTATGGAAGTGAATAATTATCATCTGAGGAAGTAGAGTGAACCCCTAGCCTTGA	2745899
QY	3954	TTTATGCTATTAATGCTGGCAATTGAGTGAATTATATCAATTACTCCGACAGATATATC	4013
Db	274588	TTTATGCTATTAATGCTGGCAATTGAGTGAATTATATCAATTACTCCGACAGATATATC	2745239
QY	4014	AGCGTTAAGCCTTATTTCTTCGTCAATTATGTTGAATGTTTCAACGCTAAGTACAAACC	4073
Db	274528	AGCGTTAAGCCTTATTTCTTCGTCAATTATGTTGAATGTTTCAACGCTAAGTACAAACC	274469
QY	4074	ACGGTAAATCTCGAGGTGTGCACAAACATTTGACGTTATGGCAAAAAGAGTGGGA	4133
Db	274468	ACGGTAAATGACAGATGTTGCACAAACATTTGAGCGTTATGGCAAAAAGAGTGGGA	2744098
QY	4134	TTAAAGCGAATAATTATCACTTTCCAAATTTCCGCTTTATTCGAAATCTCAAGTTCA	4193
Db	274408	TTAAAGCGAATAATTATCACTTTCCAACTTTCGCTTTATTCGAAATCTCAAGTTG	274349
QY	4194	CAACTGGGCAAAACGCAAAATGAGGCGCTGAAATGGGCTATCGTGGTAAATAATCA	4253
Db	274348	CAACTGGGCAAAACGCAAAATGAGGCGCTGAAATGGGCTATCGTGGTAAATAATCA	274289
QY	4254	TAAATTATCGTTATTTGATTAACAAGGTGGTCAATCAATCCACCTTTTATATCC	4313
Db	274288	TAAATTATCGTTATTTGATTAACAAGGTGG-----GCAATCCACCTTTTATATTC	2742434
QY	4314	AATTAAT 4319	
Db	274233	AATTAAT 274228	

```

RESULT 7
US-10-080-505-12
Sequence 12, Application US/10080505
Publication No. US20030073166A1
GENERAL INFORMATION:
APPLICANT: St. Geme, Joseph W.
TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINMS
FILE REFERENCE: A-59941-1/FT/DCE/DIR
CURRENT APPLICATION NUMBER: US/10/080, 505
PRIOR FILING DATE: 2002-02-22
PRIOR APPLICATION NUMBER: US 08/296, 791
PRIOR FILING DATE: 1994-10-25
PRIOR APPLICATION NUMBER: US 09/839, 996
PRIOR FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 58
SOFTWARE: Patentln version 3.1
SEQ ID NO 12
LENGTH: 5245
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: CDS
LOCATION: (430)..(4740)
OTHER INFORMATION:
US-10-080-505-12

```

```

Query Match      69.2%; Score 2990.4; DB 15; Length 5245;
Best Local Similarity 82.0%; Pred. No. 0;
Matches 3660; Conservative 0; Mismatches 636; Indels 168; Gaps 12

QY      1 TCATAGTCGTTTACTA-GATCTTTTATACGAAAAATTAATTATTAATAATTAACATT 59
          370 TCATAGTCGTTTACACGATATTTTATACGAAAAATTAATTATTAATAATAACATT 429

QY      60 ATGAAAAAAATGCTATTTCGTTCTAATTTTAAACCGTTGCATTCACTAGGGAATGTA 119

```


Db 430 ATGAAAAAAGCTGATTTTCCTGTAACCTTTTAAACCGCTTGCACTTTCATAGGAGTACTA 489
 Qy 120 TCGCAAGCTGGGCTGGTCACTTATTTTGGATGATTAACAATATATGATGATTT 179
 Db 490 TCGCAAGCTGGGCTGGTCACTTATTTTGGATGATTAACAATATATGATGATTT 549
 Qy 180 GCCGAAATTAAGGGAAGTTCAAGCTTGGGCTCAAAATATTAAGTTTATACAAACA 239
 Db 550 GCTGAAATTAAGGGAAGTTTCACTTGGGCTCAAAATATTAAGTTTATACAAACA 609
 Qy 240 GGGCAATTAAGTTGGCACTCAATGCAAAAGCCCGATGATTTATTTCTGATGTC 299
 Db 610 GGGCAATTAAGTTGGCACTCAATGCAAAAGCCCGATGATTTATTTCTGATGTC 669
 Qy 300 GGTAAAGGCTGGGCTGGTGGAAATCATATATTTGAGCGGTCATTAAGCTA 359
 Db 670 GGAATGGGCTGGGCTGGTGGAAATCATATATTTGAGCGGTCATTAAGCTA 729
 Qy 360 GGAATTAAGGCTGGTGGTGGCAAGGGAACCAACCCGATCAACATGTTTACT 419
 Db 730 GGAATTAAGGCTGGTGGTGGCAAGGGAACCAACCCGATCAACATGTTTACT 789
 Qy 420 TATAGATTTAAAGAAATTAATTAACCAAAAGATATTTACTCTTTGAGAGAT 479
 Db 790 TATCAATTTAAAGAAATTAATTAACCAAGCAAGCAATTAATCAATGATGAC 849
 Qy 480 TACATATTCAGATTAATTAATTAATGATTAACAGAGGCTCCATATGATGATG 539
 Db 850 TATCAATGCTGCTGGTGGCAAAATTTGTCAGAGCTGTAACCAAGAAATGATGAC 909
 Qy 540 AATATGAATGGCACTTATTTAGATTAACCAAAATTAACCAAGCTGTTGATGAC 599
 Db 910 AATATGAATGGCACTTATTTAGATTAACCAAAATTAACCAAGCTGTTGATGAC 969
 Qy 600 TCTGAGAGGCACTTTGGGGAATGATTAACCAAGGCAAGGCAAGGCAAGGCAAG 644
 Db 970 AAGAGTAAACATATGAGGAGATGATTAACCAAGGCAAGGCAAGGCAAGGCAAG 1029
 Qy 645 GTTCCGCTGATATATTTGATGATGATGATGATGATGATGATGATGATGATGAT 704
 Db 1030 GTTCCGCTGATATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1089
 Qy 705 AATGATTAATGATTTGGAGAGGATGTTGTAAGCGGAGATTAATGATGATGAT 764
 Db 1090 AATGATTAATGATTTGGAGAGGATGTTGTAAGCGGAGATTAATGATGATGAT 1149
 Qy 765 AATGATTAATGATTTGGAGAGGATGTTGTAAGCGGAGATTAATGATGATGAT 824
 Db 1150 AATGATTAATGATTTGGAGAGGATGTTGTAAGCGGAGATTAATGATGATGAT 1209
 Qy 825 AATGATTAATGATTTGGAGAGGATGTTGTAAGCGGAGATTAATGATGATGAT 884
 Db 1210 AATGATTAATGATTTGGAGAGGATGTTGTAAGCGGAGATTAATGATGATGAT 1269
 Qy 885 TTTCAATTTGATGATTTTATTTGATGATA---TTTCAAGAGATTTTATGATCA 941
 Db 1270 TTTCAATTTGATGATTTTATTTGATGATA---TTTCAAGAGATTTTATGATCA 1329
 Qy 942 TCACTTTAAGCCGAGCTGATTAAGAGTGAACAATTAAGGAAATGATAGTAC 1001
 Db 1330 AATTTTAAAGCCGAGCTGATTAAGAGTGAACAATTAAGGAAATGATAGTAC 1052
 Qy 1002 GGGTCTATTAATC-----AGAAATCAGAAATACCAATTAAGGAAATTAAGT 1052
 Db 1390 GGGTCTATTAATC-----AGAAATCAGAAATACCAATTAAGGAAATTAAGT 1449
 Qy 1053 GGAATATGATTTAGCTTTGAAAGAGAGATTAAGTCAATATCTTGAATGAGCA 1112
 Db 1450 GGAATATGATTTAGCTTTGAAAGAGAGATTAAGTCAATATCTTGAATGAGCA 1509
 Qy 1113 CCAATATTTATTTCTCAAGCTTTAAACATGAGGAAAGCTATATTTTATGATGATA 1172
 Db 1510 GGAATATGATTTAGCTTTGAAAGAGAGATTAAGTCAATATCTTGAATGAGCA 1569

Qy 1173 CAAGATCATTAATCTTGGATGATGATTAACCAAGGCGGCTGATTTATTTGAG 1232
 Db 1570 AAGGATCTTTAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAAT 1629
 Qy 1233 GATTAATTTAAGATTTCTCAATTTCTAACAACCTTGGCAAGAGCTGATCAATGTA 1292
 Db 1630 GATTAATTTAAGATTTCTCAATTTCTAACAACCTTGGCAAGAGCTGATCAATGTA 1689
 Qy 1293 AGTAAATATGACCGTTACTTGAAGAGTAAATGGGTTGAAATGATGATGATTTCTAA 1352
 Db 1692 GATTAATATGACCGTTACTTGAAGAGTAAATGGGTTGAAATGATGATGATTTCTAA 1749
 Qy 1353 ATTGTAAGGATCAATTTGACGCTTGAAGCAAGGCAAGGCAAGGCAAGGCAAGG 1412
 Db 1750 ATTGTAAGGATCAATTTGACGCTTGAAGCAAGGCAAGGCAAGGCAAGGCAAGG 1809
 Qy 1413 GGGATGTAAGGATCAATTTGACGCTTGAAGCAAGGCAAGGCAAGGCAAGGCAAGG 1472
 Db 1810 GGGATGTAAGGATCAATTTGACGCTTGAAGCAAGGCAAGGCAAGGCAAGGCAAGG 1869
 Qy 1473 AGTAAATTTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1532
 Db 1870 AGTAAATTTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1929
 Qy 1533 GATACGATTAATTTTATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1592
 Db 1930 GATACGATTAATTTTATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1989
 Qy 1593 TTAACCTTTAAAGCTATTAACCAATTAACCAAGGCAAGGCAAGGCAAGGCAAGG 1652
 Db 1990 TTAACCTTTAAAGCTATTAACCAATTAACCAAGGCAAGGCAAGGCAAGGCAAGG 2049
 Qy 1653 ACTCAAGCCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1709
 Db 2050 ACTCAAGCCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2209
 Qy 1710 AATTAATTTAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1769
 Db 2110 AATTAATTTAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2169
 Qy 1770 AAGATTAATTAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1829
 Db 2170 AAGATTAATTAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2229
 Qy 1830 GATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1889
 Db 2230 GATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2289
 Qy 1890 AATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1949
 Db 2290 AATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2299
 Qy 1950 TCGAAATTAAGGATTAACCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 2009
 Db 2344 TCGAAATTAAGGATTAACCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 2403
 Qy 2010 AATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2069
 Db 2404 AATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2463
 Qy 2070 TCTTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2129
 Db 2464 TCTTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2523
 Qy 2130 CCAATTAATTAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2189
 Db 2524 CCAATTAATTAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2583
 Qy 2190 AAGTGAATTTAAGGATTAACCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 2249
 Db 2584 AAGTGAATTTAAGGATTAACCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 2643

QY	2250	TCATATTAATTAACGTAATATGACAGCGCAATGTTAAAGTTAGCAAACTTAATAGCG	2300
Db	2644	TCATATTAATTAACGTAATATGACAGCGTAATTCATAGTTTAGCAAACTTAATAGCG	2700
QY	2310	AATGTCACCTTAAACCAATACAGCCCAATTTACATTAGCAACAATGCAACCCCAATAGCG	2360
Db	2704	AATGTCACCTTAAATTAATCAATAGCCCAATTTACATTAGCAACAATGCAACCCCAATAGCG	2760
QY	2370	AATATTCGACTTTCGACAACTTAACTGCAACGCTGATTAATGCAAACTTGAACGGTAAT	2420
Db	2764	AATATTCGACTTTCGACAACTTAACTGCAACGCTGATTAATGCAAACTTGAACGGTAAT	2820
QY	2430	GTGCAATTAAACGAAATTCAGCTCAATTTTCTTTAAAAAACAACCAATTTTCGACCAAAAT	2480
Db	2824	GTGCAATTAAACGAAATTCGCTCAATTTTCTTTAAAAAACAACCAATTTTCGACCAAAAT	2880
QY	2490	CAGGAGACAAAGGACACACAGCTGACGTTGGAAAAATGCGACTTGACAATGCTAGCGAT	2540
Db	2884	CAGGAGGACAAAGACACACAGTACGCTGGAAAAATGCGACTTGACAATGCTAGCGAT	2940
QY	2550	ACTACATTCGACAAATTTAACGCTAAATACACACAGTACGATACGTTAAATTCAGCTTAATCA	2600
Db	2944	GCCACATTCGACAAATTTAACGCTAAATATATGACCTGTTAAATTCAGCTTAATCA	3000
QY	2610	GCTAGCTCAAAACAATAGCCACGT---CGCGCTCAATAGAGCGGAAACAAGCCACACA	2660
Db	3004	GCTAGCTCAAAATATAGCCACGTACACCGCGCTCAATAGAGCGGAAACAAGCCACACA	3060
QY	2667	TCGGCAACAATCGTTTCCACACATTGACAGTAAATGTGTAATCTGAGTGGCAGACACA	2726
Db	3064	TCGGCAACAACATCGTTTCCACACATTGACAGTAAATGTGTAATCTGAGGCGGACAGACACA	3123
QY	2727	TTCCAAATTAATCTTATCTTATTTGGCTAATAAAGCATAAATTTAAATATTCATCATGAC	2786
Db	3124	TTCCAAATTAATCTTATCTTATTTGGCTAATAAAGCATAAATTTAAATATTCATCATGAC	3183
QY	2787	GCTAGAGGCGANTACATATATCTGTGCGAACAAGGCAAAAGAACCCGAAACCTCTGAG	2846
Db	3184	GCTAGAGGCGANTACACATATCTGTGTGCGAACAAGGCAAAAGAACCCGAAAGCCTTGAG	3243
QY	2847	CAATTAATCTTGTGTTGAAGCAAAAGATATCAACGTTATCAGTAAGTCGAAATTTACT	2906
Db	3244	CAATTAATCTTGTGTTGAAGCAAAAGATATTAACGTTATCAGCAAACTCAATTTTACT	3303
QY	2907	TTAGAAATGACACGTTGATGACAGTGCATTAACGTAAATTAATGAGAAATGATGCG	2966
Db	3304	TTAGAAATGACACAGTTGATGACAGTGCATTAACGTAAATTAATGAGAAATATGCG	3363
QY	2967	GAATTCGCTGCTACATCCCAATTAAGACAGAGAAATGGCAATGATTTGTAAGACA	3026
Db	3364	GAATTCGCTGCTACATCCCAATTAAGACAGAGAAATGGCAATGATTTGTAAGACA	3423
QY	3027	GAGCAAGCAAGAACCAATTTAGAGCCCAACAAGTTGAACCGACTGCTTAAGACAACAACA	3086
Db	3424	GAGCAAGCAAGAACCAATTTAGAGCCCAACAAGTTGAACCGACTGTAAGACAACAACA	3483
QY	3087	GGTAGCCCAAAAGTGGCGTCAAGAAAGCACGCG-----	3159
Db	3484	AGTAGTAGCAAGTGGCGGTCAAAAAGAGCGGTGTTCTGTATCCCTGCTGATCAAAAC	3543
QY	3120	-----	3119
Db	3544	CAGTTAGAGTATTAACAAGCCCAACAAGTTGAACCGACTGMAAAACAAAATAATAG	3603
QY	3120	-----AGAGCAGGCTTTCCTGATACCGCTGCTGATCAAAACCGT	3158
Db	3604	GCATAAAAAAGTCGCGTCAAAAAGAGCGGTGTTCTGTATCCCTGCTGATCAAAACCG	3663
QY	3159	TTAAACGCAATTAGAGCCCAACAAGCTGAATCAGCTGTGAAACAACAATAAGTAAGGCA	3218
Db	3664	TTAGAGTATTAACAGCCCAACAAGTTGAACCGACTGCTGMAAAACAAAATAATAGCA	3723
QY	3219	AAAACAAAAAAGTGGCGTCAAAAAGAGCGGTGTTCTGTATCCCTGCTGATCAAAAC	3278

3724 AAAAAAGTGGGTCAAAAAGACGGCAGAGAGCTTTTGTGATACCCGCTTATCTTAAGC 3783

3729 CTGTTCC---GATTAGAACCGCAGCTTGAAGTATTTAGATGCCCAACAGCAATGGCAAAA 3335

3784 CGGTAAAGGATTAGAGATCAAACTTGAGGTATTTATATGCCCAACGACAGTGAAAAA 3843

3336 GATGCTCTAGCTCAAGAAAGACGGCAAAAAACAACGGCAACAAAAAGACTGATCACCGT 3395

3844 GAACTCAAGTCA-----GAGAAACAACGGCAACAAAAAGACTGATCACCGT 3894

3396 TATTCAAATATGCTGTATACAGATTTATCTTCAACGTAATAGTATGCTTCTGTCTCA 3455

3895 TATTCAAATATGCTGTATACAGATTTATCTTCAACGTAATAGTATGCTTCTGTCTCA 3954

3456 GATGAATATAGTCTTTTGTGTAATCAGCAACAATCTGCCGTGTGGAATAATTCGA 3515

3955 GATGAATATAGTCTTTTGTGTAATCAGCAACAATCTGCCGTGTGGAATAATTCGA 4014

3516 CAGATTAATAAGACGTATATCTGATGCGTCCGTATTACAGACAGCAAAAGAAC 3575

4015 CAGATTAATAAGACGTATATCTGATGCGTCCGTATTACAGACAGCAAAAGAAC 4071

3576 TTACGTCAATTTGGGTGCAAAAAGCCTTACTGATGACGAAATTTGGGCGATTTCTCG 3635

4072 TTACGTCAAAATTTGGGTGCAAAAAGCCTTACTGATGACGAAATTTGGGCGATTTCTCG 4121

3636 CATAGCCGTCAATATACCTTGTATGAACAGGTATAAATAACAGCATTTAAACGATG 3695

4132 CATAGCCGTCAATATACCTTGTATGAACAGGTATAAATAACAGCATTTAAACGATG 4191

3696 ATGTGCGCTTTTCCCAATATCAATTTGGGCGATTTTCAATTTGGTGTAAACGTGGAAACG 3755

4192 ATGTGCGCTTTTCCCAATATCAATTTGGGCGATTTTCAATTTGGTGTAAACGTGGAAACG 4251

3756 GGAATACGTGCGGTAAATTTGGGTGAAGAAACAAACCGCAAAATTTCTGCAAAACGAT 3815

4252 GGAATACGTGCGGTAAATTTGGGTGAAGAAACAAACCGCAAAATTTCTGCAAAACGAT 4311

3816 AATTATGCGGTGATGCAAGTTATCAGTTCCGTTAGGGCAATTTGGGCAATTCAGCCTAT 3875

4312 AATTATGCGGTGATGCAAGTTATCAGTTCCGTTAGGGCAATTTGGGCAATTCAGCCTAT 4371

3876 TTGGAGTTAACTGCTATTTTATGAAACGTGAATTTCAATCTGACGAGTGAAGT 3935

4372 TTGGAGTTAACTGCTATTTTATGAAACGTGAATTTCAATCTGACGAGTGAAGT 4431

3936 AAAAGCGCTGCTGCAATTTATCGCTAATATGCTGCAATTCAGTGTATTAACATTT 3995

4432 AAAAGCGCTGCTGCAATTTATCGCTAATATGCTGCAATTCAGTGTATTAACATTT 4491

4433 AAAAGCGCTGCTGCAATTTATCGCTAATATGCTGCAATTCAGTGTATTAACATTT 4491

3996 ACTCGACAGATTAATCAACGCTTAACGCTTATTTCTTGTGTAAATATGCTGATTTTCA 4055

4492 ACTCGACAGATTAATCAACGCTTAACGCTTATTTCTTGTGTAAATATGCTGATTTTCA 4551

4056 AAGCTTAAGCTAATAACGACGCTTAATCTCAACGCTGCTGCAACACCATTTGAGAGTAT 4115

4552 AAGCTTAAGCTAATAACGACGCTTAATCTCAACGCTGCTGCAACACCATTTGAGAGTAT 4611

4116 TGGCAAAAAGATGGGATTAAGGACGAAATTTTACATTTCCAAAATTTCCGCTTTATC 4175

4612 TGGCAAAAAGATGGGATTAAGGACGAAATTTTACATTTCCAAAATTTCCGCTTTATC 4671

4176 TCAAAATCTTAAGCTTCACACTGCGGCAACACGCAAAATGCGGCGGTGAATTTGGGCTAT 4235

4672 TCAAAATCTTAAGCTTCACACTGCGGCAACACGCAAAATGCGGCGGTGAATTTGGGCTAT 4731

4236 CGTTGTAAATAATCAACATATTTATTCGTTATTTGATTAACAAGAGTGGGTAGATTCGA 4295

4732 CGTTGTAAATAATCAACATATTTGATTCGTTATTTGATTAACAAGAGTGGGCGAGATC--- 4788

4296 TCCACCTTTTATTCCAATAAT 4319

```

Db      4789 --CCACCCTTTTATTCATATAT 481C
RESULT 8
US-10-080-505-16
; Sequence 16, Application US/10080505
; Publication No. US20030073166A1
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RPT/DCF/DHR
; CURRENT APPLICATION NUMBER: US/10/080,505
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 4828
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; PEPTIDE:
; NAME/KEY: CDS
; LOCATION: (313)..(4548)
; OTHER INFORMATION:
US-10-080-505-16

Query Match      68.8%; Score 2970; DB 15; Length 4828;
Best Local Similarity 82.3%; Pred. No. 0;
Matches 358; Conservative 0; Mismatches 645; Indels 124; Gaps 10;

QY      1 TCAATAGTGGTTAACTA-GTATTTTATACGAAATTAATTAATAATACAT 59
Db      253 TCAATAGTGGTTAACTA-CGATTTTAAACGAAATTAATTAATAATACAT 312
QY      60 ATGAAAAAAGTATTTTCTTAAATTTTAAACCGCTTCATTTAGGATATG 119
Db      313 ATGAAAAAAGTATTTTCTTAAATTTTAAACCGCTTCATTTAGGATATG 372
QY      120 TCGAAGGTGGGGGCTGCTACATTTTGGGANTGATTAACAAATTAATGATTT 179
Db      373 TCGAAGGTGGGGGCTGCTACATTTTGGGANTGATTAACAAATTAATGATTT 432
QY      180 GCGAGATTAAGGGAAGTTCACAGTTGGGGCTCAAAATTAAGTTTATACAA 239
Db      433 GCGAGATTAAGGGAAGTTCACAGTTGGGGCTCAAAATTAAGTTTATACAA 492
QY      240 GGGCATTTGTTGGCATCATGATGACAAAAGCCCGATGATTTTCTGATGTC 299
Db      493 GGAATTTGTTGGCATCATGATGACAAAAGCCCGATGATTTTCTGATGTC 552
QY      300 GGTAGGCGGTGGGAGCGCTTGTGAAATCAATATATGCTGAGGTGGCATTA 359
Db      553 GGAATGGGCTGGGAGCGCTTGTGAGGTGAGTATGAGTGGCATTAATGTA 612
QY      360 GATATATACAGATTTGATTTGGTGCAGAGGAAACAACCCGATCAACATGTT 419
Db      613 GCGATATACAGATTTGATTTGGTGCAGAGGAAACAACCCGATCAACATGTT 672
QY      420 TATTAAGTTTAAACGAAATTAACTAACAAAAGATTAATTAATCTTATGAG 479
Db      673 TATTAAGTTTAAACGAAATTAACTAACAAAAGATTAATTAATCTTATGAG 732
QY      480 TACCATATCCAGATTAACATTAATCTGTTACAGAGCGCTCAATTTGATGAT 539
Db      733 TACCATACCCAGATTTACATTAATTTGTTAGAGAGCGACCCCAATGATATG 792
QY      540 AATATGATGGCAGTACTTATTCATATGAAACAAATATCCAAAGCTGTTGAT 599
Db      793 GATATGAGGGAACAAATATACATGATGAGAGAAATATCCGAAAGGCTGAT 852

QY      600 TCGAGACCGAGTTTTGGCAATGATCAAGACAAAGCGAAGTTGCCGTGATAT 659
Db      853 TCGGAGTGGAGTTTGGCAATGATCAAGACAAAGCGAAGTTGCCGTGATAT 912
QY      660 CATATCTGACAGCTGGCAATACACATATCAGCTGAGAGGATATGATAT 719
Db      913 CATATCTGACAGAGCAATACACATATCAGAGGAGGAGGAGGAGCTGATAG 972
QY      720 TTGGAGGCGATGTTGTAAGCGGAGAAATATGTCATTAACGATTTGAGCT 779
Db      973 CTGAGCGCGATGTTGTAAGCGGAGAAATATGTCATTAACGATTTGAGCT 1032
QY      780 GGGGACAGTGGTTCTCCGATTTTATTTATGATGCTGAAAAACAAATATG 839
Db      1033 GCGGACAGCGTTTGGCTATTTTATTTATGAGCGAAAAACAAATATGAT 1092
QY      840 GGAATATTAAGGAAAGGCAACCTTTGAAAGCAAGAAATGAGTTCAATG 899
Db      1093 GCGGATTTAGAGAACCGGCAACCTTTGGGCGGAGACAGAAATACATTC 1152
QY      900 AATCTTATTTTGTAAATTTTGAAGAGATTAATACATCACTTACACCGAG 959
Db      1153 AAGCTTTTGTATTAATCTTGAAGAAATTTGCGATCTGTTATGACCTAT 1212
QY      960 GGTAAATGAGGTATCAATATGATGAAATGATATGTCAGGGGTCTATAC 1019
Db      1213 GCGATATGTCATACACATTAACAGACAAAGGAGGAGGAGGATTTGACAA 1272
QY      1020 TCGAATATCCATCAAAATTAATTAAGTATGCAATATGATTTACTTGA 1079
Db      1273 ACGAGAAACATCTATAGCTGATGCGATGCTTTAAAGCAATATTA 1332
QY      1080 AAGATTAAGTATCAATCTATGATATGACGACCTAATTTATCTCAGGTT 1139
Db      1333 AAGGCAATGTTTACCA-----ATACCAAGGTCTCAATATATCTGCTG 1386
QY      1140 AATGAGAAAGCTATATTTATATGATCAAAACCAAGATCATTAATCT 1189
Db      1387 AAGGTGGAAGCTGATTTTGGAGATCAAAACCGGACGTTACTTACAA 1446
QY      1200 ATTAACTAAAGGCGGCTGCTTATTTTGAAGGTAATTTTACAGATTT 1259
Db      1447 ATCAACCAAGGTGCGGCTGCTTATTTTGAAGGTAATTTTACAGATTT 1506
QY      1260 AACCAACTTGGCAAGAGCTGCAATGATGAGTAAATTAACAGCTTACAA 1319
Db      1507 AATGCAATTTGGCAAGAGCTGCTGATGATGATGATGATGATGATG 1566
QY      1320 GTAAATGAGCTGATGATGATGATGATGATGATGATGATGATGATG 1379
Db      1567 GTAAATGAGCTGATGATGATGATGATGATGATGATGATGATGATG 1626
QY      1380 GCGAAAGGAGAAATTAAGTTCATGAGTATGAGTATGAGTATGAG 1439
Db      1627 GCGAAAGGAGAAATTAAGTTCATGAGTATGAGTATGAGTATGAG 1686
QY      1440 CAGGCAAGCATCAAGGCAACAAAGAGCTTATGTAATTTGCTTGTG 1499
Db      1687 CAGGCAAGCATCAAGGCAACAAAGAGCTTATGTAATTTGCTTGTG 1746
QY      1500 GGGATCTGTAATTAACATGATTAACATTTGATCCGATTAATTTAT 1559
Db      1747 GGTATGCTGTAATTAACATGATTAACATTTGATCCGATTAATTTAT 1806
QY      1550 GGTGATGCTGATATCTTAAGGAGCTTATTAACGTTTAAAGTATCA 1619
Db      1807 GGTGATGCTGATATCTTAAGGAGCTTATTAACGTTTAAAGTATCA 1866
QY      1620 GACGAGGAGGCAATGATTTGAAACATTAATCAACCTCAAGCGCTA 1679
Db      1867 GATGAGGAGGCAATGATTTGAAACATTAATCAACCTCAAGCGCTA 1926
QY      1680 GGAACGAAAGCATTTGTTACTTAATGAAATATTAATTAATTTA 1739

```

Ds	1927	GGCAGGATACCATTAATGACAACCTGGGATTTAACAATAAAGGAT-----	1977
Qy	1740	GAATTCCTACACGGTGGTTTGGCGAAACGATTAATAACATACATGCGCATTA	1799
Ds	1978	---ATTGCTTTAATGGTGTGGTTGGTGAATAAGATGATCTAAATATCTGAGCGTTTG	2034
Qy	1800	AACCTATTATTAACCAACACAGAGATCGTACCTTGACCTTCAGGTGGTAACAT	1859
Ds	2035	AAATGTTACTACATCCGCTTACCAAAATACATCTCTCTACACAGTGGAAACAT	2094
Qy	1860	TTAAAGCGGATATTCCCAACAAAGGTAACTATTTCACGGTGAACGACACCG	1919
Ds	2095	TTAAAGGCAATATTACTCAACACGGTGGACCTTAGTGTGTAGTGTGGCCACACCA	2154
Qy	1920	CAGCCCTAACATCTTAAATTAAGTGGTGCAGAAATGAAAGGTATACCAAGCGAA	1979
Ds	2155	CAGCAATACATCAATTAAA-----TGGCTTAAACGAGCTTGGCGACTTAAAGCGAA	2208
Qy	1980	ATTGTGGGATCCAGATTGGATCAACCGTACATTTAAAGCTGAACCTCCAAATAAA	2039
Ds	2209	GTGGTTATTGATGACGATTGGATCAACCGTACATTTAAAGCTGAACCTCCAAATAAA	2266
Qy	2040	GGCGGAATGGGGGTGGTTTCCGCAATGTTCTTCAATTGAGGAAATTTGACAGTGAC	2099
Ds	2269	GGCGGAATGGGGGTGGTTTCCGCAATGTTCTTCAATTGAGGAAATTTGACAGTGAC	2328
Qy	2160	AATATGCAAAATGCACAATGGTGTGTGTCCAAAATCAACAAATATCACTTGGACGGT	2159
Ds	2339	AATTAAGCCAAACGACATTTGGTGTGTGTCCAAAATCAACAAATATCACTTGGACGGT	2388
Qy	2160	TCAGATTGGAACGATTAACGACTTGCAAAAAGTGAATTAACGATTAACAAATTAAT	2219
Ds	2389	TCAGATTGGAACGATTAACGACTTGCAAAAAGTGAATTAACGATTAACAAATTAAT	2448
Qy	2220	AATTTATACCAAAAACAAATCAATAGCTCATTAATTTAACTGATTAAGCAAGCGC	2279
Ds	2449	GATTCATACCGACACACAAATTAATAGCTCATTAATTTAACTGATTAAGCAAGCGC	2508
Qy	2280	AATGTTAAGTTTATGCAAACTTAAATGCGAATGTCATTTAAACAATCAAGCCAAAT	2339
Ds	2509	AATGTTAAGTTTATGCAAACTTAAATGCGAATGTCATTTAAACAATCAAGCCAAAT	2568
Qy	2340	ACATTAGCAACATATGCAACCAATAGCGAATTTGCACTTTGCGACAAATTCATGCA	2399
Ds	2569	ACATTAGCAACATATGCAACCAATAGCGAATTTGCACTTTGCGACAAATTCATGCA	2628
Qy	2400	ACGGTGAATATGCAAACTTGAACGGTATATGTCATTTAAGGATTCAGCTCAATTTCT	2459
Ds	2629	ACGGTGAATATGCAAACTTGAACGGTATATGTCATTTAAGGATTCAGCTCAATTTCT	2688
Qy	2463	TTAAAAAAGCACTTTTGGACCAAAATTCAGGAGACAAAGGACCAACATGACGTTG	2519
Ds	2689	TTAAAAAAGCACTTTTGGACCAAAATTCAGGAGACAAAGGACCAACATGACGTTG	2748
Qy	2535	GAATATGCACTTGGACCAATGCTTAGCGATATCACTTGAAGAAATTTAAAGCTTAATAC	2579
Ds	2749	GAATATGCACTTGGACCAATGCTTAGCGATATCACTTGAAGAAATTTAAAGCTTAATAT	2808
Qy	2586	AGTAGCATCAGTTAAATTCAGCTTATTCAGCTAGCTCAACATTAAGCCACGTT--CGC	2636
Ds	2869	AGTAGCATCAGTTAAATTCAGCTTATTCAGCTAGCTCAACATTAAGCTAGCCACGCT	2868
Qy	2637	CGTTCAATTAGAGACGAACCAACGCAACATGCGGAAACATGCTTCAACATTTGACA	2686
Ds	2869	CGTTCAATTAGAGACGAACCAACGCAACATGCGGAAACATGCTTCAACATTTGACA	2928
Qy	2697	GTAAATGTAAATTTAGATGGGACAAAGGACATTCGATTAATCTTCACTTTATTTGGCTAT	2756
Ds	2929	GTAAATGTAAATTTAGATGGGACAAAGGACATTCGATTAATCTTCACTTTATTTGGCTAT	2988
Qy	2757	AAAAGGATTAATTTAAATTTCAATGACGCTGAGGGCGATTAATTAATTAATCTGTTGCG	2816

Dh	2989	AAAGGCAATATAATATAATTATCCAAATGACGCTGAAGGCAATTACACATTAAGCTGTCC	3048
Oy	2817	AAACAAGGCAAAAGACCCGCAAAACCTTTAGCAATTAACCTTGGTTGAAGCAAAATAT	2876
Dh	3049	GACACAGGGAAGAAGACCTGTGACCTTGGACAAATTAACCTTATTAAGAGCTGTGCAATAT	3108
Oy	2877	CAACCGTTATCAGATAGAGTCAAAATTTACTTATGAAATAACACAGTTGATGCAAGTGA	2916
Dh	3109	CACCTTTGCCAGATAGCTATAAAATTAACCTTTAAAAATTAACAGCTTGATGCGGTGGA	3168
Oy	2937	TTACGTTTAAATTAAGTGAAGATGATGCGCAATCCGCTTGCATTAACCAATAAAGAG	2996
Dh	3169	TGGCGTTATGATTAAGTGAAGAAAAAGCGCAATTCGCTGTGATATCAATTAAGAG	3228
Oy	2997	CAGGATTTGCAATGATTTAGTAAAGCAAGCAAGCAAGCAATTTGAAGCCAA	3356
Dh	3229	CAGAAATTTGCCAATGATTTAGTAAAGCAAGCAAGTAAAGCAATTTGAAGCCAA	3288
Oy	3057	CAAGTTGAACCGATGCTGTAAACAACAACAGGTGAGCCAAAGTGGGTCAAGAGAGA	3116
Dh	3289	CAGCTGAACCTGATCTAATAAAACAAAACCTGAGCTTAAGTGGGTCAAA	3342
Oy	3117	GCAAGAGCAGCGTTTCCATTAACCTGCTGATCAAGCTGTTAAAGCAATTAAGACC	3176
Dh	3343	---AGAGGGGGTTCATATACCCCGCTGATCAAGCCAGTTAAAGCAATTCAAGCG	3399
Oy	3177	AAACAAGCTGAACCTGATCTGTGAAGCA---AAAGTAAAGCAAAAACAAAAGAG	3231
Dh	3400	GAACTGAGACGATTAATATCCCAACAGATGAGCAACAGCGTCAAAATCAAGAACTA	3459
Oy	3232	TGGCGTCAAAAAGAGA---GTGTTTCTGATCCCTGCTTATGATCAAGCGCTGTGCA	3287
Dh	3460	ACTGCACTTAACCAAAAAGAACGACAGATTAACAATCAATCAATTAAGCAAAATTAATGTC	3519
Oy	3288	TTTGAAGCCGCACTGAGGTTATTAAGCCCAAC	3321
Dh	3520	TTGGCACTGATTTGGTGAAGAAAAGAACCGCTCAGATTAATTTGCTAATGCAAAATTA	3579
Oy	3332	-----AGCATCGGA	3332
Dh	3580	GCTCAGTTGATTTAACAACAACAACACTGAAGAAAAGCTTACAGATGCTGACAGCAAGCA	3639
Oy	3333	AAAGATCGTTAGCTCAAAAGAAAGCGGAAAACAAGCAACAAAGAACTTGATACG	3392
Dh	3640	AAAGACCGTAAAGCTCAAGAGCAAGCAAAAGACAACGCAAAAGAACTTGATACG	3499
Oy	3393	CGTTATCAAAATAGCGTTATCAGATTAATCGCAACAGTAAATAGATGCTTCTGT	3452
Dh	3700	CGTTATCAAAATAGCGTTATCAGATTAATCGCAACAGTAAATAGATGCTTCTGT	3759
Oy	3453	CAGATGAATTAAGTCGTCTTTTGTAGATCAAGCAATCTGCGTGTGCAAAATATC	3512
Dh	3760	CAAGATGAATTAAGTCGTCTTTTGTAGATCAAGCTCAATCTGCGTGTGCAAAATATC	3819
Oy	3513	GCAACAGATTAAGAAAGCAAGTATGATTCGATGCTTCCGTCTTATCAACACAGAAAGG	3572
Dh	3820	TCAACAGATTAAGAAAGCAAGTATGATTCGATGCTTCCGTCTTAT---CAACAAAGAG	3876
Oy	3573	AACTTAACGTCAATTTGGGAGTGAAGAAAGCTTACTATGACGAATTTGGGCAATTTTC	3632
Dh	3877	AACTTCGTCAATTTGGGAGTGAAGAAAGCTTACTATGACGAAGATTTGGGCAATTTTC	3916
Oy	3633	TGCAATAGCGCTCAATTAATCTTGTAGTAAAGAGTTAAATCAAGCAAGTAAAG	3692
Dh	3937	TGCAATAGCGCTCAATTAATCTTGTAGTAAAGAGTAAATCAAGCAAGTAAAG	3996
Oy	3693	ATGATGTGGGTTTGGCCAAATCAATGGGCGAATTTACATTTGTTGTTAAAGCTGGGA	3752
Dh	3997	ATGATGTGGGTTTGGCCAAATCAATGGGCGAATTTACATTTGTTGTTAAAGCTGGGA	4056
Oy	3753	AGGGAATCAAGTCGAGTAAATGGCTGAAGAAACAAGCCGAAATTAATCATGAAAGCG	3812
Dh	4057	ACGGGAATTAATGCGATTAATGGCTGAAGAAACAAGCCGAAATTAATCATGAAAGCG	4116

Db 1522 ACTTTACATTTGAAATTAATTAATCAAGTGTGAGTGTGATTTGTAAGTAAAT 1587-
Qy 1239 TTTTACAGATATCCAAAT---TCTAACCAACTTGGCAAGAGCTGGATACATGTAAGT 1295
Db 1582 TTTGTTGTAAGGCAATCAAAATATATATACCTTGGCAAGTGTGAGGTTCTGTGGA 1641
Qy 1296 GAAATATGACCTGTTACTTGTGAAAGTAATGAGCTGGAAATGATGCACTTTCTAAAT 1355
Db 1642 GAAGAAATGCTGTATATAGGAGGTGATATTCAGAAAGGCGATCGCTATTCAAAAT 1701
Qy 1356 GGTAAAGGACATTCACCTTCAGCCGCAAGGAGGAAATTAAGGTGATCAAGTAAAGC 1415
Db 1702 GGGCTGGGACCTTACTTGTATGTATGAAAGGAAACCTTAAAGAGCTGATGTGCT 1761
Qy 1416 GATGTAAAGTCACTTTTGGAGCAGAGGCAAGATCAAGGCAAAACAAAGCCTTAA 1475
Db 1762 AACGTTTGTGTGTAGTATCAACAGCAGATGATTCAGGTCAAAACAAAGCCTTAA 1821
Qy 1476 GAAATGTGCTGTGTTAGCGGCAAGAGGACTGTTCAATTAACATGATTAACAAATTA 1535
Db 1822 GAAGTGGCATGTATAGTGTAGAGCTACGTTCAACTAAATAGTGCATCAAGTGTAT 1881
Qy 1536 ACCGATTAATTTATTTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1595
Db 1882 CCTAACATATTTATTTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1941
Qy 1596 ACCTTAAACGATCCAAATATCGAGCAGAGGAGGCAATGATTTGAACCAATATCACT 1655
Db 1942 ACCTTGAAGGATCCAAATATCGAGTGAAGGAGGAGGATTTGAACCAACGCTTCT 2001
Qy 1656 CAAGCCCTATGCTCTATTTCTGAGGAAGAAAGCATTTGTTCTAATTAAGAAATAT 1715
Db 2002 CAAGCCGCAATATTCAGATTAAGGCAAGCAACTATTAATTA-----GAT 2049
Qy 1716 ATTAATTAACCTGATTAACGAAAGAAATTTGCCCAACGCTGTGTGTGTGTGTGTGT 1775
Db 2056 AGCAAAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2109
Qy 1776 AAAATTAACAAATGAGGCAATTAACCTTATTAATTAATTAATTAATTAATTAATTA 1835
Db 2116 AAAGCTAAACAAATGAGTGTGTAAATGTATTAATTAATTAATTAATTAATTAATTA 2169
Qy 1836 TTGCTACTTCAAGTGTGTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1895
Db 2170 TTGTTGCTTCTGAGGAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2229
Qy 1896 TTTTCAAGGCTAGACCGGCAACGCGACGCTCAATCAATTAATTAATTAATTAATTAAT 1955
Db 2230 GTTTTATGT 2289
Qy 1956 ATGAAAGGATTAACCAAGAGGAAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2015
Db 2290 ATGAAAGGATTAACCAAGAGGAAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2349
Qy 2016 AAAGCTAAACCTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2075
Db 2350 AAAGCTAAACCTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2409
Qy 2076 ATGAGGGAATTTGACAGTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 2135
Db 2410 ATGAGGGAATTTGACAGTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 2469
Qy 2136 CAACAAATTAACATTTGACGCGTTCAGATTTGAGCAAGATTAATTAATTAATTAATTA 2195
Db 2470 CAGCAAAATTAACATTTGACGCGTTCAGATTTGAGCAAGATTAATTAATTAATTAATTA 2529
Qy 2196 GATTAACGATTAACAAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2255
Db 2530 GATTAACGATTAACAAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2589
Qy 2256 AATTAACGATTAACAAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2315
Db 2590 AATTAACGATTAACAAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2649

Qy 2316 ACTTAACAAATTAACAGGCAATTTAATTAACCAACATTAATTAATTAATTAATTAATTA 2375
Db 2650 ACTTAATTAATTAACAGGCAATTTAATTAACCAACATTAATTAATTAATTAATTAATTA 2709
Qy 2316 GAGCTTTCGCAATTAACCTGCAAGGTGTATTAATTAATTAATTAATTAATTAATTAATTA 2435
Db 2710 AAACCTTCAATTAACAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2769
Qy 2436 TTAACGATTAACGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2495
Db 2770 TTAATGATTAACGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2829
Qy 2496 GACAAAGGCAACAGTGTGAGGAAATTTGCAATTAATTAATTAATTAATTAATTAATTAATTA 2555
Db 2830 GGGAAAGGCAACAGTGTGAGGAAATTTGCAATTAATTAATTAATTAATTAATTAATTAATTA 2889
Qy 2556 TTGCAAGATTAACGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2615
Db 2890 TTGCAAGATTAACGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2949
Qy 2616 TCAGAAATTAACGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2669
Db 2950 TCAGAAATTAACGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3009
Qy 2670 GCAAGACATGCTTTCACACATTTGACATTAATTAATTAATTAATTAATTAATTAATTAATTA 2725
Db 3010 GCAAGACATGCTTTCACACATTTGACATTAATTAATTAATTAATTAATTAATTAATTAATTA 3069
Qy 2730 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2789
Db 3070 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3129
Qy 2730 GAGGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2849
Db 3130 GAGGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3189
Qy 2850 TTAATTTGTTGAAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2909
Db 3190 TTAATTTGTTGAAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3249
Qy 2910 GAAATATGACAGCTTGAATGAGGTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 2969
Db 3250 GAAATATGACAGCTTGAATGAGGTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 3309
Qy 2970 TTCCGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3029
Db 3310 TTCCGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3369
Qy 3030 CAAGCAGACGATTAAGAGGCAACAGTGAACGCTGTAAACCAACAGT 3089
Db 3370 CAAGCAGACGATTAAGAGGCAACAGTGAACGCTGTAAACCAACAGT 3429
Qy 3090 GAGCCAAAGTGTGATTAAGAGGCAACAGTGAACGCTGTAAACCAACAGT 3149
Db 3430 GAGCCAAAGTGTGATTAAGAGGCAACAGTGAACGCTGTAAACCAACAGT 3480
Qy 3150 CAAGCCTGTTTAAGGATTAAGAGGCAACAGTGAACGCTGTAAACCAACAGT 3209
Db 3481 CAAGCCTGTTTAAGGATTAAGAGGCAACAGTGAACGCTGTAAACCAACAGT 3537
Qy 3210 AGTAAGGCAAAACCAAAAGGCGGTCAAAAGAGCAGTGTGTGTGTGTGTGTGTGTGTGTGT 3269
Db 3538 AGTAAGGCAAAACCAAAAGGCGGTCAAAAGAGCAGTGTGTGTGTGTGTGTGTGTGTGTGT 3597
Qy 3270 GATCAAGCTGT 3329
Db 3598 GATCAAGCTGT 3648
Qy 3330 GAAAAAGATGCTGT 3386
Db 3649 AAAAAAGATGCTGT 3708

```

1 CURRENT APPLICATION NUMBER: US/10/080,505
2 CURRENT FILING DATE: 2002-02-22
3 PRIOR APPLICATION NUMBER: US 08/296,791
4 PRIOR FILING DATE: 1994-10-25
5 PRIOR APPLICATION NUMBER: US 09/639,996
6 PRIOR FILING DATE: 2001-04-20
7 NUMBER OF SEQ ID NOS: 58
8 SOFTWARE: PatentIn version 3.1
9 SEQ ID NO 10
10 LENGTH: 4605
11 TYPE: DNA
12 ORGANISM: Haemophilus influenzae
13 FEATURE:
14 NAME/KEY: CDS
15 LOCATION: (422)..(4597)
16 OTHER INFORMATION:
17 US-10-080-505-10
18
19 Query Match          62.7%; Score 2739.6; DB 15; Length 4605;
20 Best Local Similarity 79.4%; Pred. No. 0;
21 Matches 3401; Conservative 0; Mismatches 809; Indels 76; Gaps 13
22
23      1 TCAATGTCGTCTACTA-GTATTTTTAADCGAAANATTCTVATTAATAAACATT 59
24      362 TCAATAGTCGTITACACGGATTTTTTAATAGAAAATACTVATTAATAATACATT 421
25
26      63 ATGAAAAAACTGATTTGCCTCTATTTTTPACGGCTTGCAATTCATTAGGAATAGA 119
27      422 ATGAAAAAACTGATTTGCCTCTATTTTTPACGGCTTGCAATTCATTAGGAATAGA 481
28
29      120 TCAGCAGCGTGGCTGGTACAATTATTTGGGATGTATACCAATATATTCGATTTT 179
30      482 TCAGCAGCGTGGCTGGTACAATTATTTGGGATGTATACCAATATATTCGATTTT 541
31
32      180 GCGSAAATAAAGGGAAGTTCAGGTTGGGCTCAAAATTTAAGTTTAAACAAACA 239
33      542 GCGSAAATAAAGGGAAGTTCAGGTTGGGCTCAAGATATTAATTCACAAATAAAAA 601
34
35      240 GGGCAATTAGTTGCCATCATCATGCAAAAGCCCGATGATGATTTTTCGTAGTGTCA 295
36      602 GGGGAAATGATAGGTACGATGATGAAAGGTGTGCTATGCTGATTTATCTCCATGTT 661
37
38      300 CGTACAGCGGTGGTAGCTTGGTTGAAAATCAAATATGTGAGGTGGCACATACGTA 359
39      662 CGGTGTGTATTCACATTGTATAGTAGACGACATTTAATTAGCTGGCACATATGTA 721
40
41      360 GGAATATACAGATGTGATTTTGTGACAGGGGAAACACCCTGATCAATCATCGTTTACT 419
42      722 GGGTATGATGTCTTATTTTGTATGTAGAGGGGAAATTCASACCAATCGTTTAAAG 781
43
44      420 TAATAGATTGAAACGAATTAAGTACAAAAGAATTAATTATCATCTTATAGAGAGAT 479
45      782 TAATAGATTGAAACGAATTAATTAATTAAGACGGGTATGACATATPAT-----GAT 835
46
47      480 TACCAATAACAGATTAATATAATTTGTTAAGAGGGCTCCGATTAATAGACTTGG 539
48      836 TATCAACATCCAGATTTAAGAAATTTGTACAGGAATCTGCATTAATGAAATGGTTTCA 895
49
50      540 AATATGAATGGCAATCTTATTCAGATAGACAAATATTCAGAAAGTGTTCGATAGCGC 539
51      896 TATATGAAATGGTATCATTAACAAATTTTATTAATCAATTCCTTTCGAGTTAGATGSA 955
52
53      600 TCTGACACGGCAGTTTGGGAAATGATCAAGCAAAAGC---GACCAAGTTGCCGTGCA 656
54      956 AGTGGCAATCAATGGTGGAAAGACGAATATATTAACCAATTTGAGACATTAAGCCTTGA 1015
55
56      657 TATCATATTCGACAGCTGGCAATACACATCAAGCTGGAGCAGTAATGATATTCG 716
57      1016 CGTTCATGCTTAATAGTGGAAATTCCTTTGAAGATGAGACAGCTGGTATAGGTATATA 1075
58
59      717 TATTGGAGGCGCATTTCTGTAAGCGGGAATATGTGTCAT--ACGATATGAGAGCTCA 776
60      1076 GAATTAATATGGGAGATCAAAATCTCAATTAATATATGTGCTCATCTACAGCGAGCTTA 1135

```

QY 777 AAGGGGACAGTGGTCTCCGATGTTTATGATGCTGAAAAACAAAATGTTAAT 836
 Db 1136 TTGGGGATAGTGTCTCCAAATGTTATGATGATGAAGATTAAGAAATGGTTATTA 1195
 QY 837 AATGGATATTAAGGGAAGCAACCTTTTGAAGCAAAAGAAATGGCTTCAATGGTT 896
 Db 1136 AATGGCTGTACGGGAAGAAATCTTAATGCTGATGGAACAGCTATCAATTAACA 1255
 QY 897 CGCAATCTTATTTGATGATAATTTGGAAGAGATTACATCATCCTTACACCGGA 956
 Db 1256 GAAAAAATTTATTTCAAGATTTCTTAATCAAGACATTAACGCTAATTTTGGATTC 1315
 QY 957 GCTGTAAATGAGTGTACAAATTAAGTGAATGATTAATGCTCAAGGCTC----- 1006
 Db 1316 AATGCTAATATAGATTTAATATAGGAGTGAACCAATGGAAGATGCAACATGAAA 1375
 QY 1007 --TATACCTCAAAATCCGGAATTCATCAAGAAATTAATAATACCTTACAAATTAAGT 1064
 Db 1376 AGTACATTAATCTAAAAAGCTTAATCAGCTGAACGAATGAGGCTCTTATGATTAATAGC 1435
 QY 1065 TTACCTTTGAAAGAAAGATTAAT-----GTCTAATTCCTGATATAGCGAATTAAT 1118
 Db 1436 CAATCTCATGATGCTAGATTAAGATTAAGATGCGATGAATTCCTCTTATTAAGTCTTAAT 1495
 QY 1119 ATTATTTCTCAAGTTTAAACATGGAAGAAAGCTAATTTATATGATCAAAAACAAAGA 1178
 Db 1496 CAAATGCTCCAGCATTAATCATGGAAGAAATTAATTTACCTTGGCATCAAGAAACAGGA 1555
 QY 1179 TCATTATCTTGCACTGACATTAACGAAGGAGGCTGTCTTATTTGAGGATTAAT 1238
 Db 1556 ACTTAAATTAATGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1615
 QY 1239 TTACAGATATCTCAAAAT--TCTAACAATCTTGGCAAGAGCTGCAATCATGTAAGT 1295
 Db 1616 TTGTGTGTAAGGGAATCAAAATTAATTAATTTGCAAGATGCTGAGGCTTTCTGTGGA 1675
 QY 1296 GAAATATGCAACGTTACTTGGAAAGTAAATGCGATGACATGATGATCTTAAAT 1355
 Db 1676 GAAAGAAATGATGTTGATGATGCGATGATCAATCAAGAGGATGCTAATCAAAATTT 1735
 QY 1356 GGTAAAGAACTATGCAAGCTTCAAGCAAGGGAATTAAGTTTCAATCAAGGATGAGC 1415
 Db 1736 GGGCTGGAACCTTAATGTTAATGTTAAGGGAATTAAGGAACTTAAGAAAGCTTAATGCTGAT 1795
 QY 1416 GATGTAAGTCAATTTGGAGCAGCAGCAAGCATCAAGGCAACAAACAAGCTTTAGT 1475
 Db 1796 AAGGTTGCTGTGTTAATGATCAAGCAGATGATCAGTCAAAAACAGCTTTAA 1855
 QY 1476 GAAATGCTGTGTAGCGGAGAGGAGCTGTTCAATTAACGATGATTAACATTTGAT 1535
 Db 1856 GAAGTTGGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1915
 QY 1536 ACCGATTAATTTTATTCGAGCTTGTGCTGTGCTGCTTACATCTTAAGGCAATCTTA 1595
 Db 1916 CTTAAATATTTATTTTGTGCTTGTGCTGTGCTTACATCTTAAGGCAATCTTA 1975
 QY 1596 ACCTTTAAAGTATCCAAATTAAGCAGAGGAGGCAATGATGATGATCAATTAACAT 1655
 Db 1976 ACCTTTGAAGTATCCAAATTAAGCAGAGGAGGCAATGATGATGATCAATTAACAT 2035
 QY 1656 CAAGCGCTAATGCTAATTTACTGGAACGAAGCAATGTTCTAATGAAATTAAT 1715
 Db 2036 CAAGCGCAATTAATTTACTGGAACGAAGCAATGTTCTAATGAAATTAAT 2083
 QY 1716 ATTAATTAATCTGATTAAGAAAGAAATTCCTAAGAGGTTGCTGTGCTGGAACGAT 1775
 Db 2084 AGCAATCAATCTTAATTAAGAAATTAAGATGATTAAGGCTGTGCTGTGGAACGAT 2143
 QY 1776 AAAAATTAACCAATGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1835
 Db 2144 AAAGCTAAACCAATGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2203

QY 1836 TTGCTACTTCAAGTGTGACAAATTTAAAGCGATATTACCCAAACAAAGTAACTA 1895
 Db 2204 TTGTGCTTTCTGGGGGCAAAATTTAAAGCGAATATCACCCAAAATGCTGTAGCTTA 2263
 QY 1896 TTTTTCAGGCTAGACGACACCGGACCTTCAATCATTTTAAATTAACGTTGCTAGAA 1955
 Db 2264 GTTTTATGAGTGTCTTCAAGCGCTCATCTTACATCAATTAATTAAGAGAGATCTGCTAAC 2323
 QY 1956 ATGAAGGATTAACCAAGGCGAAATGTGTGCTATCAAGATTAATTAATTAATTAAT 2015
 Db 2324 ATGAAGGATTAACCAAGGCGAAATGTGTGCTATCAAGATTAATTAATTAATTAAT 2383
 QY 2016 AAAGCTAAATCTTCAATTTAAAGCGAGTGTGCTGTGCTTCTGCAATGTTCTTCA 2075
 Db 2384 AAAGCTAAATCTTCAATTTAAAGCGAGTGTGCTGTGCTTCTGCAATGTTCTTCA 2443
 QY 2076 ATTGAGGAAATGAGACGTCAGCAATTAATGCAATGCAATTTGCTGTGCTCAAT 2135
 Db 2444 ATTGAGGAAATGAGACGTCAGCAATTAATGCAATGCAATTTGCTGTGCTCAAT 2503
 QY 2136 CAACAAATTAATCTTGTGCAAGGCTTCAATTTGCAAGATTAATGCAATTTGCAAAAGT 2195
 Db 2504 CAACAAATTAATCTTGTGCAAGGCTTCAATTTGCAAGATTAATGCAATTTGCAAAAGT 2563
 QY 2196 GATTTAACGATTAACAAAGTATTAATTTCTAATCAAAACAAATCAATGCTCTAAT 2255
 Db 2564 GATTTAACGATTAACAAAGTATTAATTTCTAATCAAAACAAATCAATGCTCTAAT 2623
 QY 2256 AATTTAATCTAATTAATGCAAGGCAATTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT 2315
 Db 2624 AATTTAATCTAATTAATGCAAGGCAATTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT 2683
 QY 2316 ACTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2375
 Db 2684 ACTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2743
 QY 2376 CAACTTCCGACATTAATCAATGCAAGGCTGATTAATGCAATTTGCAATTTGCAAT 2435
 Db 2744 AAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2803
 QY 2436 TTAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2495
 Db 2804 TTAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2863
 QY 2496 GAAAGGCAACAGTGTGCAAGGCTTGAAGGCAATGCAATGCAATGCAATGCAATGCAAT 2555
 Db 2864 GAAAGGCAACAGTGTGCAAGGCTTGAAGGCAATGCAATGCAATGCAATGCAATGCAAT 2923
 QY 2556 TTGCAATTTAATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2615
 Db 2924 TTGCAATTTAATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2983
 QY 2616 TGAAGATTAATGCA-----GTGCGCTTCAATTAAGAGGAAACAAAGGCAATG 2669
 Db 2984 TGAAGATTAATGCAAGGCTTGAAGGCAATTAATTAATTAATTAATTAATTAATTAAT 3043
 QY 2670 GGAAGGATGCTTTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2729
 Db 3044 GGAAGGATGCTTTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3103
 QY 2730 CAATTAATCTTATTTATTTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2789
 Db 3104 CAATTAATCTTATTTATTTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3163
 QY 2790 GAGGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2849
 Db 3164 GAGGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3223
 QY 2850 TTAATCTTGTGTAAGGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2909
 Db 3224 TTAATCTTGTGTAAGGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3283
 QY 2910 GAAATATCAAGCTGATGCAAGTGTGCTTAATTAATTAATTAATTAATTAATTAATTAAT 2969

81 CTTAATTTTAAACCGCTTCATTCATTAAGGATATGATCGACAGCGTGGCTGGTAC 140
73 CTTCTACTTACCATATGCTGCTGGCTTCCTCCAGACTGGCCGGGAC 132
141 ACTATTTTGGATTTGATTCACATATATCGTATTTTCCGAGATTTAAAGGAACTTC 200
133 ACTATTTTGGATTTGATTCACATATATCGTATTTTCCGAGATTTAAAGGAACTTC 192
201 ACAGTTGGGGCTCAAAATATTAAGTTTAAACAAACAGGCGATTTAGTTGGACATCA 260
193 GCACTGGGGCCAAAGATTTGAGTTTCAACAAAAGGAGTTGGTGGCAAAATCA 252
261 ATGCAAAAGCCCGGATGATTTTCTGATGTCAGTAAACGCGCTGGCAGCTTG 320
253 ATGCAAAAGCCCGATGATTTTCTGATGTCAGTAAACGCGCTGGCAGCTTG 312
321 GTTGAATTCATATATTTGAGCGTGGACATTAACGTAAGATATTCAGATTTT 380
313 GTGGGATTCATATATTTGAGCGTGGACATTAACGCGCTGGCAGCTTG 372
381 GTTGAAGGAGGAAACACCCGATTCACATGCTTTTCTTATTAAGTTGTAACGAAT 440
373 GGTGGAGAGAGAGAAATCCCGATTCAGACGCTTTTCTTACCAATTTGTAAGGAAT 432
441 AACTCAAAAAGATATTTTCACTTATGAGACATTAACGATTAACGATTAACAT 500
433 AATATTAAGCTGCAATTCACACCTTTCAACGCGGATTAACGATTAACGCTTTGCA 492
501 AATTTGCTTACGAGAGCGCTCCATTTGATTAAGCTTCAATTAAGGAGATTAAC 560
493 AATTTGCTTACGAGAGCGCTCCATTTGATTAAGCTTCAATTAAGGAGATTAAC 552
561 TCAGATAGCAAAATATTCAGACGCTTTGATGCTTGGACGCGGCTTTGGCG 620
553 TCCGATTAAGAGAAATATCCCGATTCAGACGCTTTTCTTACCAATTTGTAAGGAAT 612
621 AATGATCAAGCAAGGCGACCAAGTGGCGGATTAATTTATTTGACAGCTGGCAT 680
613 TATGATGATGCAAAACAGCGGCTTATCTTACCGGCGCATTTGATTTGGGGGCAAT 672
681 AOCACATACGCGTGGACGAGATTAAGATTTGATTTTGGAGCGGATTTGTA 740
673 AOCATATGCAAGGCTGGGAGAAATATGCGTATGTTGAGCGGATTTGTA 732
741 GCGGAGATATGCTCATTAACGATTTGAGGCTCAAGGCGGAGATTTGCGATG 800
733 GCCAAGCATATGCGCTATGCGGATTTGAGGCGGAGGAGCAAGCGCTTGGCAATG 792
801 TTTATTTATGATGCTGAAAAACAAATGCTTAATTAATGAGATTAACGGAAGCAAC 860
793 TTTATTTATGACAAAACAAATTAATGCTGCTCAACGAGATTTTCAACCGGCTAC 852
861 CTTTGAAGGAGAAATGCTTTGATTTGTTGGAATTTATTT---TGATGA 917
853 CTTATTTCCGAGGAGAAAGGCTTTCCAGCTGATGCAAGATTTGCTTACAGTAC 912
918 AITTTGAAAGATTTTACATATCACTTACACCGGAGTGTATGAGTATCA 977
913 AITTAAGAGGATTAACATACGCTTTTGAACCGGAGATTAACGAGCATTTTTC 972
978 AITTAAGGAGATTAATGCTGAGGCTTTTAACTGAAATTCAGATTAACATCA 1037
973 TTTATTCACAAACAAACGCTTACGCTTACGAGTAAACGAAACGAAAGCTTCAAT 1032
1038 AITTAATTTAGGTTAGCAATTAATGAGTTT-----ACCTTGAAGAGAGATTA 1088
1033 CCAAGCTTTAAAGTACAGACGCTGCTTTTGAAGATTTTGAAGAACTGATTA 1092
1089 GTTCAATATCTAGATTAACGAGCTTATATTTTCTCACTTTTAAACATGAGAA 1148
1093 GAACGATTTAGGCGAGGCGGCTTATACGATACGCTCAAGGTTAAACAGCTGA 1152
1145 AGCTATATTTTATGATCAAAACAAAGATTAATTTCTGCACTGACATTAACCA 1208

1153 AACCTTCTTTATTCATTAACGCAACGCAAACTCATTTATCAACCAATCAACCA 1212
1209 GGGGCGGCTGCTTTATTTTGAAGGATTAATTTACGATATCTCCAAATTTCAACCAAT 1268
1213 GGGGCGGCTGCTTTATTTTGAAGGATTAATTTTGAAGGATTTTGAAGGATTTTGAAG 1272
1269 TGGCAAGAGCTGGCATATCATGATTAAGAAATAGCAAGCTTACTGGAAGATTAAG 1328
1273 TGGCAAGAGCTGGCATATCATGATTAAGAAATAGCAAGCTTACTGGAAGATTAAG 1332
1329 GTGAAACATGATGACCTTTCTAAATTTGTAAGAAACATTTGCAAGCTTCAAGCAAG 1388
1333 GTGGCAAAACAGCGCTGCTCAAAATTTGCAAAAGGAGCTGCAAGCTTCAAGCAAG 1392
1389 GAAATTAAGGTTGATTCAGGCTAGGCGATTTGTAAGATTTTGAAGCGAGCGAGAC 1446
1393 GAAACCAAGGCTGATTCAGGCTAGGCGAGGATTAAGATTTTGAATTCAGCGAGAC 1452
1449 GATCAAGGCAACAAACGCTTTAGTGAATTTGCTTGGTGGTGGAGGAGGATGTT 1508
1453 GATTAAGGCAAAACAGCTTTGTAAGATTTGCTTGGTGGTGGAGGAGGATGTT 1512
1509 CATTTAAACGATTAACAAATTTGATTAACGATTAATTTTATTTTGGCTTTGGTGGT 1568
1513 CACTGAATGCGATTAATCAATTCACCCGACAACTCTATTTGGCTTTGGCGGAG 1572
1569 CGCTTAATGCTTAACGCGATTTATTAACCTTTAAACGATTCGAAATTAAGCAAGG 1628
1573 CGTTGAATTTAAACGAGATTTGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1632
1629 GCAATGATTTGAAACATTAATTAACCTCAAGCGCTTAATGCTATTAATTTAGGAGCA 1688
1633 GCGATGATTTGCAATTAATTAACCAACCAATTCACCTTATTAAGGAGATTA 1692
1689 AGCATTTGCTTACCTTAATGAAATTAATTAATTAATTTGATTAACGAAAGATTTGCC 1748
1693 AGTATTAACACAGCGATTTGAAATTAATTAATTAATTTGATTAACGAAAGATTTGCC 1752
1749 TACAAAGGTTGTTGGCGAAACCAATTAATTAATTAATTAATTTGATTAACCTTAT 1808
1753 TACAAAGGTTGTTGGCGAAACCAATTAATTAATTAATTTGATTAACCTTAT 1812
1809 TATTAACCAACCAACAAAGATTTGCTTGGCTTACCTTACCTTACCTTACCTTACCTT 1868
1813 TACAGCGCGCGGAGAAAGCGGCTTCTGCTTGGGAGAAATTTAAACGCG 1872
1869 GATTTTACCAACCAACAAAGATTTGCTTGGCTTACCTTACCTTACCTTACCTTACCT 1928
1873 AACATCAAGCAACAAAGGCTTCTGCTTGGCTTACCTTACCTTACCTTACCTTACCT 1932
1929 AATCATTTTAATTAAGCTTGGCTGCAATTAAGGATTAACCAAGGCGAATTTGGTGG 1988
1933 AATCATTTTAAGGCGGCTGCTGCAATTAAGGATTAACCAAGGCGAATTTGGTGG 1992
1989 GATCAAGATTTGATTAACGCTTACCTTAAAGCTTAAAGCTTAAAGGCGAAGT 2048
1993 GATCAAGATTTGATTAACGCTTACCTTAAAGCTTAAAGCTTAAAGGCGGAG 2052
2049 GGGGCTGCTTGGCTGCAATTTTCTCAATTTGAGGAAATTTGAGAGCTTACATTAATG 2108
2053 GGGGCTGCTTGGCTGCAATTTTCTCAATTTGAGGAAATTTGAGAGCTTACATTAATG 2112
2109 AATGCAATTTGCTTGGCTGCAATTTTCTCAATTTGAGGAAATTTGAGAGCTTACATTA 2168
2113 CAGGATTTTGGCTGCAATTTTCTCAATTTGAGGAAATTTGAGAGCTTACATTAATG 2172
2169 ACGGATTTAAACGCTTCTCAAAAGGATTTAAACGATTAACGAAAGTATTAATTTCTTA 2228
2173 ACGGCTTGAACATTTGCTGCAAAACCAATTTACGAGATTAAGGATTTGCTTCTTAATG 2232
2229 CCAAAACCAATTTCAATGCTTATTAATTTTAACTGATTAATGCAAGGAGATTTGAA 2288

Db 2233 ACTAAGACGACATAGCGGCAATGTGACCTTGGCCATCAACGCTCATTTAAATCTCA 2292
 Qy 2289 GGTTAAGCAAAAATTATGSCATGTCACTTTAACTAACTACAGCCATTTACATTAAGC 2348
 Db 2293 GGGGTTCACACTCACTCAAGCCNAICTTACTGCAATGGCCATACAGTTATACGTGAGC 2352
 Qy 2349 AACATCCACCCCAATAGGCAATTTTCG----- 2377
 Db 2353 CACAGCCACCACCAAAACGGCAACTTAGCTCGTGGGCAATGCCAGCAACATTTAAT 2412
 Qy 2378 ----- 2377
 Db 2413 CAAGCCATTTAAACGGCAACATGCGCTTGGGCAATGCTTCATTTAATCTAAGCAAC 2472
 Qy 2378 ----- ACTTCCGCAATTCACATGSCATGCGATTAAT 2411
 Db 2473 AACCCGGCAAAAAGGCAATGTGACGCTTTCGCAACAGCTAAGCAACGTAAGCAT 2532
 Qy 2412 GCAAACTTAACGTAATGTGCAATTTACGATTCAGCTCAATTTTCTTTAAAAACAGC 2471
 Db 2533 TCCGCACTCAACGGCAATGTCTCCCTAGCCCAATAGGCAATTCATTTGAAAAACAG 2592
 Qy 2472 CATTTTGGACCAAAATTCAGGAGACAAAGSCAACAGTGAAGCTTGGAAAAATGCGACT 2531
 Db 2593 CGCTTACCGCAACACTAGCGAGCAAGATACAGCTTACACTTAAAGACAGCGCA 2652
 Qy 2532 TGACAAATGCTAGCGATTAATTAAGTTCAGAAATTTAAAGCTAAATTAACATGATCAG 2591
 Db 2653 TGAACGCTGCGCTCAGGACAGCAATTTAGCAATTTAACTTCAAGCCCACTTAC 2712
 Qy 2592 TTAAATCAGCTTAT-----TGAGCTAGCTCAACATTAAG 2627
 Db 2713 CTCAATTCGCGCTATGCGCAAGATGCTGAGGCGCGCAACCGGAGTGTGTGACAGAG 2772
 Qy 2628 CCAAGTCCGCGCTCAT-----TGAGAGCGAAACAGCCCACTGCGCAAGCAT 2678
 Db 2773 CCGGCGCGCGCTGCGCGCTTCCCTATTATCCGTACCCGCAACTTCGCTGTAATTC 2832
 Qy 2679 CGTTTCAACAATGACAGTAAATGTAATTAAGTGGGCAAGGCACTTCCATTTACT 2738
 Db 2833 CGTTTCAACGCTACAGGTAAACGGCAATTTGAACGCTTACAGAACTTCGCTTTAG 2892
 Qy 2739 TCATCTTTATTGCGCTATAAAGCGATAAATTAATTAATCAATGACGCTGAGGCGAT 2798
 Db 2893 TCGGAATCTTCGCGCTACCAAGCGCAAAATTAAGCTGCGGAAGTTCGGAAGCACT 2952
 Qy 2799 TACATATTAATCTGTGCGAACACAGGCAAAAGCCGAACCTTGAAGATTTAATCTTG 2858
 Db 2953 TACACTTGGCGGTCAACATACCGGCAACGACCCGTAAGCTCGATCAATTAAGCGTA 3012
 Qy 2859 GTTGAAGCAAGATATATCAACGTTATCAATAGCTCAATTTACTTTGAAAAATGAC 2918
 Db 3013 GTTGAAGGGAAGAACAAACCGCTGTCCGAAAACTTAATTTTCACTCTGCAAAAGAA 3072
 Qy 2919 CAGGTGATGCAAGGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2978
 Db 3073 CAGGTGATGCGGCGCTGCGCTTACCAACTCAATCCGCAAAAGCGCAAGTTCGCGCTG 3132
 Qy 2979 CATACCCCAATTAAGAGAGGAAATGCAATGATTTATTAAGGCAAGGCAAGCAAGAA 3038
 Db 3133 CATATTCGGTCAAAAGAAACAGAGCTTTCGCAAACTCGGCAAGGCAAGGCAAGAA 3192
 Qy 3039 GCAACATTAAGAACCA-----ACAACTGAACCGACTGTAAACCAACAAAGGTGACCA 3095
 Db 3193 CAGGCGGAAAAAGAACAGCGCAACAGCTTGAAGCGCTGATTCGCGCGGCGGAGTCC 3252
 Qy 3096 AAGGTGCGGTCAAGAAAGAGAGGAGAGAGGTTTCTGATACCTGCTGATCAAGC 3155
 Db 3253 GCCGAAAGAGAGAAAGGTTCGCAACCGGCGCGCGCAAGCGAGCGGAAAAATGTGCG 3312
 Qy 3156 CTGTTAAAGCATTAAGAGCAAAACAGCTGAATCTGACTGTAACCAAAAAAGTAAG 3215
 Db 3313 ATTAATGACGCGAGAGAAAGAAAAACGGTGAAGCGGAT----- 3354

Qy 3216 GCAAAAAAGAGAGGCGGTCAAAAAAGAGAGTGTTCATTCCTCTGATATCA 3275
 Db 3355 -----AAGAGAGCGGCTTGGCAAAACAGGCGCAACGCAACCGCGCGCTATCA 3406
 Qy 3276 AGCTGTTGCGATTAGAACCGCACTTGAAGTTATTAATGACCCCAAGCAATGCAAAAA 3335
 Db 3407 CGGCTTCCCG 3459
 Qy 3336 GATCGTCTAGCTCAAGAGAGAGCGGAAAAACAGCAAAACAAAGAACTGATGACGCT 3395
 Db 3460 CAACCGCAACTCAACCCCAACCG-----CAGCGGAGCTGATCAACGCT 3504
 Qy 3396 TATTCAAATGATGCTATCAAGATTTATGCAACAGTAATTAATTAATTAATTAATTA 3455
 Db 3505 TATTCCAATGATGCTATGATGATTTATTCGCAACGCTCAACAGGTTTTCGCGTACG 3564
 Qy 3456 GATGATTAATGATGCTTTTGTGATCAAGCAAACTGCGGTGTAACAAATATGCA 3515
 Db 3565 GACGATTAAGAGCGCGCTGTTTTCGCAAGACCGCGCAACGCGTTTTCGCAAGCGCATC 3624
 Qy 3516 CAGGATTAAGAGCGCTATGATTTGATGCTTTCGCTTATCAGCAGCAAAAGCAAC 3575
 Db 3625 CGGACACCAACATACCGTTGCAAGATTTTCGCGCTTACCGCAAC---AACCGAC 3681
 Qy 3576 TTAAGTCAATTAAGGAGTGAACAAACCTTAGCTTAAGCAAGATTAAGGAGGAGTTTCG 3635
 Db 3682 CTGCGCAATCGTATGCAAGAAACCTGCGAGCGGCGGCTGCGCATCTGTTTTCG 3741
 Qy 3636 CATACCGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3695
 Db 3742 CACAAACCGAACCAAAACCTTTCGCAAGACGCGATGCGCACTGCGACGCGTTCGCA 3801
 Qy 3696 ATGTGCGTTTTCGCGCATATCATATGCGGCGATTTCAACTTGTGTAAACGTGGAGAG 3755
 Db 3802 GCGCGCGTTTTCGCGCATATGCGCATGCGAGGTTGCACTTGGCATACAGCGGCGCG 3861
 Qy 3756 GGAATCAGTGGAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3815
 Db 3862 GGTTTTACAGCGCGAGCTTTCAGACGCGATCGAGGCAAAATCCCGCGCGCTG 3921
 Qy 3816 AATTAAGCGTAATCAAGTATACGTTCCGTTTACGCAATTAAGGCAATTAAGCTTAT 3875
 Db 3922 CATACCGCATCAAGCAAGTACCGCGCGGTTTCGCGGATTCGCGCATGCAACGCTAC 3981
 Qy 3876 TTTGAGTTAATGCTATTTTATGACGTAATTAATTAATTAATTAATTAATTAATTA 3935
 Db 3982 ATGCGCGCAAGCGCTATTTGTTCCAAAGCGAATTAACCGCTACGAAACGTCATATC 4041
 Qy 3936 AAAACGCTTACGCTTATTAATGCTATTAATGCTGCAATTCAGTTGATTAATTAAT 3995
 Db 4042 GCGACCGCGGCTGTTGCTTCAACGCTACCGGCGGATTAAGGAGATTAATTAATTA 4101
 Qy 3996 ACTCCAGATTAATTAATGAGGTTTACGCTTATTTTCGCTTATTAATTAATTAATTA 4055
 Db 4102 AAACCGGCGCAACATTTCAATCAACGCTTATTAATTAATTAATTAATTAATTAAT 4161
 Qy 4056 AAGCTTAAGTAACCAACGCTTAATTCAGGTTGTAAGCAACCAATTTGAGCGTTAT 4115
 Db 4162 TGGGCAAGGTCGAAACGCTCAATTCGCGCTTATGAGATTTGCGCAAAAC 4221
 Qy 4116 TGGCAAAAGAGTGGGATTAAGGCAAGAAATTTTCAATTTCCGCTTTTATC 4175
 Db 4222 CGGATGCGGATGAGGCGGTAAAGCGCAATTAAGGTTTCAAGCTGCTCCCAACGCT 4281
 Qy 4176 TCAAAATTCAGGTTCAATTCGCGCAACAGCAAAATGCGGCGTAATTAATTAATTA 4235
 Db 4282 GCGCGCGCAAAAGCGCGCACTGGAAGCGCAACAGCGCGGCGATCAATTAATTAATTA 4341
 Qy 4236 CGTTGCTAA 4244
 Db 4342 CGCTGTTAA 4350

QY	1444	AGTCATTTGGAGACGAGCGACAGCATCAAGGCAACAAACAAGCCTTAGTAAATTGG	14383
DB	1455	AGTCATCTTAGTACGACGCGGAGATCAAGCAAAAACACCTTTAGTAAATCGG	15144
QY	1484	CTTGSTTACCGCAGAGGGACGTTCATTAAACGATGAATAACATTGATACCGATA	15433
DB	1515	CTTGGTCACCGCAGGGGGACGTTCACTGAATGCGATATACGTTCAACCCGACAA	15734
QY	1544	ATTCTATTTGGCCTTTGTTGCTGATGATCTTAAGCGGATCTTAATTAACCTTAA	16033
DB	1575	ACTCTATTTGGCTTTGCGCGGACGTTGGATTTGAACGGCATATGCTTGTGTCCA	16344
QY	1604	ACGATTCACAAATAACGACACGAGGGGCAATGATTTGAACCATTAATCAACTCAAGCCG	16633
DB	1635	CCGATTCACAAATAACGATGAAGGGGCGATGATGTCAACCAATCAAGACAAAGATC	16944
QY	1664	TAAATGCACTATTACTGGGACGAAAGCATTTGTTCACTTAATGAATAATTAATA	17233
DB	1695	CACGTTACCATTAACGAGCAATPAA-----GATTTACTTCAACCGGCAATPACAACA	17483
QY	1724	ACTGATTCACAGAAAAGAAATGGCTCAACAGGTTGGTTGGCGAAACGATPAAATTA	17813
DB	1749	CTTGAATACCAAAAAGAAATGGCTCAACGTTGTTGTTGGCAAAAGATGACACA	18083
QY	1784	ACGCAATGGGCGATTAACCTTATTTAAACCAACACAGAGATCGTACTTGTACT	18433
DB	1809	AACGAACGGGCGGCTCAATCTGAATTTACCAACCGAGAACCGATTCGACTTACTGCT	18683
QY	1844	TTCAAGGTGTCAAATTTTAAAGGCGATTTACCAACAAAGTAAACGATTTTTCAG	19033
DB	1869	TTCCGGGGGAAATTTTAAACGCAATACACGCAACAAACGCAAACTGTTTTCAG	19283
QY	1904	CGGTAGACCGACACCGCAGCCTTACATCACTTTAAATGAAGTTGGTCAGAAATGGAAG	19633
DB	1929	CGGACACCGACACCGACGCTCAATCACTTTAAGAAAGGGGTGCTAAAATGGAAG	19883
QY	1964	TATPCCACAGGCGAAATTTGTGTGGATCCAGATTGATCAACCGTACATTTAAAGCTCA	20233
DB	1989	TATCCACAGGAGAAATGTTGTGTGGACAAAGATTGATCAACCGCATTTTAAAGGGA	20483
QY	2024	AAACTTCGAATTTAAAGCGGGAAGTGGGTGTTCTCGCAATGTTCTTCATTAAGGG	20833
DB	2049	AAACTTCGAATTTAAAGCGGCGGACAAAGCGGTGTTCTCGCAATGTTCTTCATTAAGGG	21083
QY	2084	AAATGGAACGTCAACATPATGCAAAATGCCAATTTGTGTTGCCAAATCAACAA	21433
DB	2109	CGATGGCATTTAAAGCAATCAACGCCAAGCATTTTGTGTGTGCCACCGCATCAAGCA	21683
QY	2144	TACCATTTCAACGGTTCAGATTGGACAGATTTAAACGCTTTCACAAAAGGATTTAC	22033
DB	2169	CTCAATCTTACAGTTGCGACTGCAAGGCTTCAACAGATTTACCGAAAAACATTAAC	22283
QY	2204	CGATACAAAGTTTATATCTTATCCAAAAACAAATCAATGCTCTATTAATTTAAC	22633
DB	2229	CGACGATTAAGGATGCTCTCATTAAGGACACCGCATCAAGGCAATGCAACTTTC	22883
QY	2264	TGATATGCAACGCGCAATTTAAAGTTTAGAAAACTTATAGCAATGCACTTTAAC	23233
DB	2289	CGATACGCTCATTTAAATCTCAAGACTTTCACACTCAACGGCAATCTTATGTCAGG	23483
QY	2324	AAATCCACCGCA-----	23533
DB	2349	CGAGACACGCACTATACGTTAACCGGCAACGCAACCAAAACGGCAACTCACTCGT	24083
QY	2336	-----	23353
DB	2409	GGGCAATGCCCAAGCAATTTATCAAGCACTTTAAACGCGCAACATCGGCTTCGCA	24683
QY	2336	-----ATTTCATTAAGCAAAATGCAACCAATTTAGGCAATATTCGACTTTCGCA	23863
DB	2469	CAATGCTCATTTATCTAAGCAACACCGCGTAAACAAACGGCAAGCTGACGCTTTCGCA	25283

2387	CATTCAACATGCAACGGTGGATTAATGCAAACTTGAAACGGTAAATGCAATTAACGATTC	2446
2529	CAGCGCTAAGGCAGAAAGTAAGCAATTCGCGATCAACGGCAATGTCCTCCAGCCGATTA	2598
2447	ACTGTAATTTCTTTAAAAAACGCCATTTTTCGACCAAAATTCAGGAGACAAAGCAC	2506
2583	GGCAGCTATTCATTATTGAAAAGACCGCTTTACCGGAAAAATTCAGCGCGCGAGAGATAC	2648
2507	AACAGTGAAGTGGAAAAAGGCACTTGACAAATGCTACAGCATCTACATTCAGATTT	2566
2643	GGCATTAACCTTAAAAAGACAGGAAATGGAGCGTCGCGCTCGGGCAAGGAAATTC	2708
2567	AAAGCTAAATTAACAGTAGATCAAGTTAAATTCAGCTAT-----	2606
2703	AAAGCTTGAACAGCCACATTAACCTCAATTCGCGCTATGACAGAGATGCGGACGAC	2768
2607	----TGAGTAGCTCAAAACAAATAGCCGACG-----TGCGGCTCATTAAGACGGA	2653
2769	GCMAACCGGACGAGCGGCGAGATGGCGCGCGCGCGCGCTTGGCGCGCTTCCCTATATCGT	2828
2654	AACACAGCCAAATCGGCGAGAACATCGCTTCAACACATTAACAGTAATGTAATGAG	2713
2829	TAGCGCGCAACTTGGGAGAAATCCGCTTCAACAGCGTGAACGGTAACGGCAAAATGGA	2889
2714	TGGGCAAGGCACTTCGCAATTTACTTCATCTTTATTTGGCTATTAAGGCAATTAATGAA	2773
2889	CGGTCAGGAAACATTCGCTTATGTGAAACCTTCGCTACCGCAGCGGCAAAATGAA	2948
2774	ATTATCCAAATGACGCTGAAGCGCAATTCATATATCTGTGCACACAGGCAAGAAC	2833
2949	GCTGCGGAAAGTTCGAAAGGCACTTACCTTGGCTGTCAACATTCGCGCAAGAAC	3008
2834	CGAAACCCCTTAGCAATTAACCTTGGTTGAAGCAAGTAATTAACCGTTATGATTA	2893
3009	CGTAAAGCTGAGCAATTTAGCGTAGTGAAGGAAAAAGCAACACACCGCTGCGAAAA	3068
2894	GCTCAATTTACTTTAGAAAAATGACACAGTTATGAGAGTGCATTAAGTTAAATTAAGT	2953
3069	TCTTAAATTTTACCTTGCAAAAGCAACGTCGATTCGCGCGATGAGGCTATACGCTAT	3128
2954	GAGAGATATGCGGAAATTCGCTTGCAATTAACCCAAATTAAGCAGGAAATTCACATGA	3013
3129	CGGAAAGAGCGGAGATTCGCGCTGCAATATCCGCTCAAAAGAACAGCTTTCGACAA	3188
3014	TTTACTAAGACAGACCAAGCAAGCAACATTAAGACCAATTAAGTGAACCACTGC	3073
3189	ACTGCGCAAGCGCGAGAAACAGAGCGCGCTTGACGCGCAAAAGGCACTACCTGGCG	3248
3074	TAAACACAAACAGTGAAGCCCAAAAGTGGCGGTCAAGAAAGSACAGGAGACGCTTCC	3133
3249	CAAAACACAGCGGAAAAAGACACAGCGCCAAAGCTTGAACGCTGATTTGGCGCGGCG	3308
3134	TGATACCTGCTGATCAAGGCTGTTAAACGCAATTAGAGCCCAACAGCTGAACCTAGC	3193
3309	CAATGCAACGAAAGAGCAAGAAAGGTTGCCAAACCGGCGGAGCAGGCGGAGAAAA	3368
3134	TGCTGAACACAAAAAGTAGAGCAAAAACAAAAAATGTCGCTCAAAAAGCAGTGT	3253
3369	TGCGCGCAATTATGACAGCGAGAAAGGAAAAACGGGTGAAGCGGATTAACACCGC	3428
3254	TTCTGATCCCTGCTTGATCAAAAGCTGTTCGATATAGAACCCCACTTGAGGTATTGGA	3313
3429	CTTGGCAAAACAGCGGAAGGAAACCGCGCGGTGACACCGCTTCCCGCGCGCGG	3488
3314	TGCCCCACAGCAATGGGAAAAAGTGGTCTAGCTCAAGAAAGAGCGGAAAAACAGCGAA	3373
3489	CCGCGCCCGCGGGAATTTGCCGCAACGCGACGCCCAACCGCAACCGCAGCGC--	3546
3374	ACAAAAAACCTTATGACGCGTATTAATAATGCTGCTTATCAAGTATATCTCAACAGT	3433
3547	-----GACCTATAGCGCTTATGCAATAGCGTTAGAGTAATTTTCGCGCAGCT	3559
3434	AAATAGTACCTTCTGTTCAAGATGAATTAAGTCGCTTTTGTAGATCAAGCAATAC	3453

```

Db 3600 CAAAGCGTTTGGCGGATGACGACGATTTGACCGCTTTGGCCGAAAGACCCGCCAA 3659
Qy 3494 TGCCGTGTGACAAATATGACAGATTAAGAAAGCTATGATTTGATGCTTCCTG 3553
Db 3660 CGCGTTTGGACAGGGGATCCGGGACACCAACATCACTGCGGAATTTCCGGC 3719
Qy 3554 TTATACAGACGAAAGAACTTATGCTCAATTTGGGCTGCAAAAAGCTTAACTG 3613
Db 3720 CTACCGCCAAAC--AAACCGACCTGCCCAATTCGATTTGAGAAAACCTCGAGAGCG 3776
Qy 3614 AGCAATTTGGGAGTTTCTGCGCAGCGCTTCAATTAATCTTATGATGAAGCTTAA 3673
Db 3777 GCGGCTGCGATCTGTTTGTGCAACGACGACGCGAAACCTTGAAGAGCGCTCG 3836
Qy 3674 AAATACGCGACATTAAGATGATGTCGGTTTTCGCCATATCAATGCGCGATTA 3733
Db 3897 CAATCGGACCGCTTGGCCACGCGTGGCGTTTGGGGAATAGGCAATCGAGTTTGA 3896
Qy 3734 ATTGTGTAAACGTGGGAAAGGGAATCGTGGGATTAATAGCTGAAAGAAAGCG 3793
Db 3897 CATGCGATCAGGCGGCGGCGGCTTTTGTAGCGCGACCTTTCAGAGCGCATCAGG 3956
Qy 3794 AAAATTCATCGCAAAAAGCATTAATTAATGCGGTAATGCAAGTTATCGCTTAAG 3853
Db 3957 CAATTCGCGCGCGCGCGCTGCTGCTTACGCGATTAAGGTAAGTACCGCGAGTTT 4016
Qy 3854 GCAATTTGCGATTCAGCTTATTTTGGAGTTAATGCTTATTTTAAAGTAAATTA 3913
Db 4017 CGATTTGCGATTCAGACCGGACATCGGCGCAACGCGCTATTTGCTCAAAAAGCGAT 4076
Qy 3914 TCATCTGAGAGAGTGAAGAGTGAAGAAAGCTTACCTTGCATTTAATGCTTAACT 3973
Db 4077 CCGATTCAGAAAGCTGCAATATGCGACCCCGGCGCTTGCATTAACCGTACCGCG 4136
Qy 3974 CATTCAGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4033
Db 4137 CATTAAGGAGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4196
Qy 4034 CGTCATTTATGTTATGTTTAAAGCTTAACGTAAACCAACGTTAATCTACGCTT 4093
Db 4197 CCGTCTTATACCGATGCGCGCTTCCGGCAAGTCCGAAAGCGGCTGATTAACGCG 4256
Qy 4094 GCAACAACTTTGACGTTATTTGCAAAAAGAGTGGATTAAGGCAAGAAATTTTCA 4153
Db 4257 GCGCGAGATTTGCGCAAAACCGCAGTGGGATTTGGGCGTAAACCGCAATCAAG 4316
Qy 4154 TTTCGAATTTGCGCTTTTATCTCAAAATCTCAAGTTTCACTCGGCAACAGCAAA 4213
Db 4317 TTTCAGCTGCTTCTTCAAGCTGCGCGCCCAAGGCGCGCAATTTGAAGCGCAG 4376
Qy 4214 TGTGGCGGTAAATTTGGGCTATGCTTGG 4241
Db 4377 GCGGCGCATCAATTTAGGCTACCGCTGG 4404

```

RESULT 13
US-10-282-122A-22193

; Sequence 22193, Application US/10282122A
; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Habelbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykand, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, S.

```

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-23
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,048
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22193
; LENGTH: 5085
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
US-10-282-122A-22193

Query Match 4.7%; Score 202; DB 13; Length 5085;
Best Local Similarity 60.9%; Pred. No. 1,78-36;
Matches 347; Conservative 0; Mismatches 220; Indels 3; Gaps 1;

Qy 1176 GATCATTTATCTTGGATGTGACATTAACCAAGGCGCGCTTATTTGAGGCT 1235
Db 1183 GGAACCTTACCTTAAATATATATGATCAACGCGCGGCTTGTCTTGAAGA 1242
Qy 1236 AATTTAGATGATTCGAATTC--TAACCAACTTGGCAAGAGCTGGCATCATGTA 1292
Db 1243 GATTTAGAGTTAAAGGACTTGTGATGATGACCTTGGAAAGGCTGGGTTCTGTT 1302
Qy 1293 AGTGAATAGACCGTTACTTGGAAAGTAAATGCGTGAACATGATGACTTTCTTAA 1352
Db 1303 GGTGATGAAAAACAGTAAGTGAAGATGATCAACCCGAATCTGATGTTAGCTTAA 1362
Qy 1353 ATTGTGAAGAAATGACGTGCAAGCCAAAGGGAAATTAAGTTCATGACGTGA 1412
Db 1363 ATGGGCAAGAAACATTAATTTAGAGAAAGGAAATTAAGTTGCTTAAAGTG 1422
Qy 1413 GCGCATGTAAGTCAATTTTGGAGCGAGGCAAGCAACCAAGCAACAGCTTT 1472
Db 1423 GCGCATGTAAGTCAATTTTGGAGCGAGGCAAGCAACCAAGCAATTAAGTAAAGCTTT 1482
Qy 1473 AGTGAATAGCGTTGTTAGCGGAGGACCTTTCAATTAAGATGATTAACAAATTT 1532
Db 1483 TCACAGTAGTAATGTAAGTGGTGGCTCACTGTTGATTAAGATGATTAAGCAAGTA 1542
Qy 1533 GATACGATTAATTTTATTTGGCTTGTGGTGGTGGCTTAACTTAAGGCAATCA 1592
Db 1543 GATCCAAATCCCATTTTCTTGGCTTAAAGGTGGTGAATTAAGCCAAATGCAATAT 1632
Qy 1593 TTAACTTTAAAGTATCCAAATAGCAAGGAGGCAATGTTGTAACCATATATCA 1652
Db 1603 CTGACTTTGAACATATCCGTAATATGATGATGCGCAAGATGATTAACCAATATCC 1662
Qy 1653 ACTGAAGCGCTAATGCTATTAATCTGGAAGCAAAAGCAATTTCTTCAATGGAAT 1712
Db 1663 AGCAAACTTACTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1722

```

Qy 1713 AATATATATAAAGCTGATTCAGCAAGAAAGAA 1742
 Db 1723 ACTGCATATATATAGACGACGACGATGAA 1752

RESULT 14

US-10-282-122A-29633
 ; Sequence 29633, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Karl
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282.122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 29633

LENGTH: 5322

TYPE: DNA

ORGANISM: Neisseria meningitidis

US-10-282-122A-29633

Query Match 3.8%; Score 164.4; DB 13; Length 5322;
 Best Local Similarity 55.1%; Pred. No. 1.3e-27;

Matches 368; Conservative 0; Mismatches 291; Indels 9; Gaps 2;

Qy 1156 ATTTATGATCAAAACAGATCATTAATCTTGATCGATTCATTAACCAAGGCGCGG 1215
 Db 1070 ATGTACCTTGAACCAACGCACTTTGGATTGGATCAAAACATCAACCAAGCGCGG 1129
 Qy 1216 GTGCTCTTATTTGAGGTAATTTTCAGT---ATTCGAAATTCACCAACTGCGG 1272
 Db 1136 GGGGCTGTTTTCACCAAGGCAATTCACAGTCAAAAGGTGCAAAATATGAGCATCTTGGC 1189
 Qy 1273 AAGAGCTGCATATCATGTAACTGAATAATGACCGGTACTTGGAAAGTAAATGCGCTGG 1332
 Db 1190 TAGTGACAGGATGTATGTGCGACGCAAAAGTGTGGCAAGTCAAAATCCGAA 1249
 Qy 1333 AACATGATCACTTTTAAATTTGTAAAGAAATGTGACGTTCAAGCCAAAGCGGAA 1392

Db 1250 ATGCGACAGATTCGCAAAATGCGCAAAAGCACATTTGAAATTAACGCGACAGCGCTTA 1309
 Qy 1393 ATTAAGTTTCATGACGATGAGTACGATGTAAGTATTTTGGAGCGACGACAGCATC 1452
 Db 1310 ACCAAGGCAATTAAGATGCGAGGATGAGGTTATTTGAAATCAAAACCGATCCG 1369
 Qy 1453 AAGGCAACCAACAGCCCTTGTGAAATTTGCTTGGTACGCGACAGAGCGACCTGTTCAT 1512
 Db 1370 ACAAABATTCAGGCTTCTCCCAAGTGGGATTTGACCGCGGCGGTACATTTGCTG 1429
 Qy 1513 TAAACATATTAACATTTGATACGATTAATTTTTCGCTTGGTGGTGGCT 1572
 Db 1430 TGAACAGTCCGACCAATTAACCGAATTAATCTGTATTTTCGCTTTCGTCGCGTGT 1489
 Qy 1573 TAGATCTTAAGGGGATTCATTAACCTTTAAAGATTCGCAAAATGCGACAGCGGCGCA 1632
 Db 1490 TGAATGCCAATGGCAATGACTTGAATCTTTGAAACATTCGCAAGTGTATTAAGCGCGC 1549
 Qy 1633 TGATGTGACCATTAATTAACATCAAGCGCTAAATGCTACTATTACTGAGAACGAAAGCA 1692
 Db 1550 GCAATGTCAACCAACACAGACGCGCGCTCGAATCAATCAATTAACCGCAATCATTA 1609
 Qy 1693 TTGTTCTACTTAATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1752
 Db 1610 TTACCGCTCTCT-----CAAAATCTTCTGTTTATGAAATCCAAATGATTAACATGATG 1663
 Qy 1753 ACGGTGTGTTGCGGAAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1812
 Db 1664 ACGATTAATTAATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1723
 Qy 1813 AACCAAC 1820
 Db 1724 ACAAAC 1731

RESULT 15

US-10-282-122A-29078

; Sequence 29078, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Karl
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282.122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

? PRIOR FILING DATE: 2001-02-09
 ? PRIOR APPLICATION NUMBER: 60/269,308
 ? PRIOR FILING DATE: 2001-02-16
 ? Remaining Prior Application data removed - See File Wrapper or PALM
 ? NUMBER OF SEQ ID NOS: 78614
 ? SOFTWARE: PatentIn version 3.1
 ? SEQ ID NO 29078
 ?
 ? LENGTH: 4779
 ?
 ? TYPE: DNA
 ?
 ? ORGANISM: *Neisseria gonorrhoeae*
 ?
 ? US-10-282-122A-29078

Query Match	3.6%	Score 155.2;	DB 13;	length 4779;
Best Local Similarity	56.9%;	Pred. No. 1.0e-25;		
Matches 305; Conservative	0;	Mismatches 228;	Indels 3;	Gaps 1;

Qy	1152	TAAATTTTAAAGATCAAAAACAAGATCAATTAATCTTCGCATCTGATCAATPACAAGG	1211
Db	1399	CAAAATGTTACCTTTGAAGACACAGCACTTTGGTTTGGATCAAAACATCAACAAGG	1155
Qy	1212	GCGGGTGTCTTTATTTTGAAGGTAAATTTTACGT---ATCTCCAAATTTCAACCAACT	1266
Db	1159	GCGGGGCGCTGTTTTTCAAGGCCATTAACACAGTCGTAAGCGGAAATTAATGACATCACT	1218
Qy	1269	TGGCAAGAGCTGCGCATACATGTAAAGTAAAGAAATAGCAACGGTACTTGGAAAGTAAATGCG	1322
Db	1219	TGGTTTAGGTGGGGGATTTGATGTTCGCCAGCGGAAAAAAGTCGTTTGGCAAGTCGAAAAAT	1278
Qy	1329	GTGGAACATGATCGACTTTCTAAATTTGGTAAAGACATTCGACGTTCAACCAAGG	1386
Db	1279	CCGAAATGGCGACAAATTTGCGAAAATTCGCGAAGGCACTTTGAAATTAACGCGACAGG	1338
Qy	1389	GAATAATTAAGGTGATCGCGGTAGCGGATGTAAAGTCACTTTTGGAGCAGAGCGAGC	1446
Db	1339	GTTTAACCAAGGSCAATTTAAAGTCGCGGACGCGATCTTATCTGAATTAAGAACCGAT	1396
Qy	1449	GATCAAGGCAACAAACAAGCTTTAGTGAATTTGGCTTGGTACGGCGAGAGCACTGTT	1506
Db	1399	TCAAAACAAAAGTTGAGGCTTCTCCCAAGTGGCAATTCACAGCGAACGGGTACATTG	1458
Qy	1509	CAATTAAAGATGATTAACAAATTTGATACCGAATAATTTTTCGGCTCTCGGTGTGT	1566
Db	1459	GTCCTGAACAGTCGGACCAAGATTAACCCGAAATATCTGTATTTTCGGTTTCCGGGCGGT	1518
Qy	1569	CGCTTGAATCTTAAACGGGCAATTCATTAACCTTTAAACGATCCAAATACGACGAGGCG	1626
Db	1519	CGTTTGAATGCCAATGGCAATGACTTGACTTTTGAACACATTCGCAAGTCGAGTGAAGGC	1578
Qy	1629	GCATATGATTTGAACCATTAATACATCTACAGCGCGCTAATGTCACTATTACTGGGAA	1684
Db	1579	GCGCGCATTTGCACCAACAACACACACAGCGCTCCGAAATCACTTAAACGGGTAA	1634

Search completed: June 24, 2004, 20:24:32
Job time : 1730 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: June 24, 2004, 07:38:46 ; Search time 13101 Seconds
(without alignments)
11599.790 Million cell updates/sec

Title: US-10-645-655-1

Perfect score: 4319

Sequence: 1 TCAATAGTGGTTTACTAGT.....ACCTTTTATTCATATAT 4319

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3757330 segs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA Main:*

1: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US100B_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US101A_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US101B_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US102A_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US102B_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US103A_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US103B_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US104A_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US104B_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US105A_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US105B_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US106A_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US106B_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US107A_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US107B_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US108A_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US108B_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US109A_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US109B_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US110A_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US110B_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US111A_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US111B_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US112A_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US112B_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US113A_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US113B_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US114A_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US114B_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US115A_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US115B_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US116A_COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US116B_COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US117A_COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US117B_COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US118A_COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US118B_COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US119A_COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US119B_COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US120A_COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US120B_COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US121A_COMB.seq.*
86: /cgn2_6/ptodata/2/pna/US121B_COMB.seq.*
87: /cgn2_6/ptodata/2/pna/US122A_COMB.seq.*
88: /cgn2_6/ptodata/2/pna/US122B_COMB.seq.*
89: /cgn2_6/ptodata/2/pna/US123A_COMB.seq.*
90: /cgn2_6/ptodata/2/pna/US123B_COMB.seq.*
91: /cgn2_6/ptodata/2/pna/US124A_COMB.seq.*
92: /cgn2_6/ptodata/2/pna/US124B_COMB.seq.*
93: /cgn2_6/ptodata/2/pna/US125A_COMB.seq.*
94: /cgn2_6/ptodata/2/pna/US125B_COMB.seq.*
95: /cgn2_6/ptodata/2/pna/US126A_COMB.seq.*
96: /cgn2_6/ptodata/2/pna/US126B_COMB.seq.*
97: /cgn2_6/ptodata/2/pna/US127A_COMB.seq.*
98: /cgn2_6/ptodata/2/pna/US127B_COMB.seq.*
99: /cgn2_6/ptodata/2/pna/US128A_COMB.seq.*
100: /cgn2_6/ptodata/2/pna/US128B_COMB.seq.*
101: /cgn2_6/ptodata/2/pna/US129A_COMB.seq.*
102: /cgn2_6/ptodata/2/pna/US129B_COMB.seq.*
103: /cgn2_6/ptodata/2/pna/US130A_COMB.seq.*
104: /cgn2_6/ptodata/2/pna/US130B_COMB.seq.*
105: /cgn2_6/ptodata/2/pna/US131A_COMB.seq.*
106: /cgn2_6/ptodata/2/pna/US131B_COMB.seq.*
107: /cgn2_6/ptodata/2/pna/US132A_COMB.seq.*
108: /cgn2_6/ptodata/2/pna/US132B_COMB.seq.*
109: /cgn2_6/ptodata/2/pna/US133A_COMB.seq.*
110: /cgn2_6/ptodata/2/pna/US133B_COMB.seq.*
111: /cgn2_6/ptodata/2/pna/US134A_COMB.seq.*
112: /cgn2_6/ptodata/2/pna/US134B_COMB.seq.*
113: /cgn2_6/ptodata/2/pna/US135A_COMB.seq.*
114: /cgn2_6/ptodata/2/pna/US135B_COMB.seq.*
115: /cgn2_6/ptodata/2/pna/US136A_COMB.seq.*
116: /cgn2_6/ptodata/2/pna/US136B_COMB.seq.*
117: /cgn2_6/ptodata/2/pna/US137A_COMB.seq.*
118: /cgn2_6/ptodata/2/pna/US137B_COMB.seq.*
119: /cgn2_6/ptodata/2/pna/US138A_COMB.seq.*
120: /cgn2_6/ptodata/2/pna/US138B_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

```

: FEATURE:
: NAME/KEY: CDS
: LOCATION: (60)..(4241)
: OTHER INFORMATION:
PCT:0503-05226-1

Query Match      100.0%; Score 4319; DB 1; Length 4319;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TCATATGCTGTTTACTACTATTTTAAATAGAAAAATTCTTAATTAAATTAACATTG 60
Db      1 TCATATGCTGTTTACTACTATTTTAAATAGAAAAATTCTTAATTAAATTAACATTG 60

QY      61 TGAATAAAACGTATTTTGGTCTTAATTTTAAACCGTGTGATTTCACTAGAGATAGT 120
Db      61 TGAATAAAACGTATTTTGGTCTTAATTTTAAACCGTGTGATTTCACTAGAGATAGT 120

QY      121 CGCAAGCTGGGCTGTGTCACTATTATTGGATTTAGATTAACAAATTAATTCGATTTTG 180
Db      121 CGCAAGCTGGGCTGTGTCACTATTATTGGATTTAGATTAACAAATTAATTCGATTTTG 180

QY      181 CCGAGATATAAAGGAAAGTCAAGTGTGGGCTCAAAATATAAGGTTATATACAAACAG 240
Db      181 CCGAGATATAAAGGAAAGTCAAGTGTGGGCTCAAAATATAAGGTTATATACAAACAG 240

QY      241 GGCATATTGTTGTCACATCATATGACAAAAAGCCCGCATGATTTTCTGTAGTGTCA 300
Db      241 GGCATATTGTTGTCACATCATATGACAAAAAGCCCGCATGATTTTCTGTAGTGTCA 300

QY      301 GTTACGGGATGACAGCTTGTTGTAATAATCAATATATGTGACGGGACATTAAGTAG 360
Db      301 GTTACGGGATGACAGCTTGTTGTAATAATCAATATATGTGACGGGACATTAAGTAG 360

QY      361 GATATACAGATGTTGATTTTGGTGCACAGGGAAACAACCCCGATCAACATCGTTTACTT 420
Db      361 GATATACAGATGTTGATTTTGGTGCACAGGGAAACAACCCCGATCAACATCGTTTACTT 420

QY      421 ATTAGTTGTATTAACAAATAATCATCAAAAAGATTAATTCATCCCTTATAGAGAGATT 480
Db      421 ATTAGTTGTATTAACAAATAATCATCAAAAAGATTAATTCATCCCTTATAGAGAGATT 480

QY      481 ACCATATTCACGATTAACATTAATTCGTTACAGAGCGGCTCAATGTAGATGATTCGA 540
Db      481 ACCATATTCACGATTAACATTAATTCGTTACAGAGCGGCTCAATGTAGATGATTCGA 540

QY      541 AATATGAATGGACGACTTTTTCAGATATGAACAAATATACAGAGCGTGTGCTATCGGCT 600
Db      541 AATATGAATGGACGACTTTTTCAGATATGAACAAATATACAGAGCGTGTGCTATCGGCT 600

```



```

Db 3121 GAGGAGGTTCCGATACCTGCGCATCAAGCCTGTAAAGCATTTGAAGCCAAAC 3180
Qy 3181 AAGCTGAACGACCTGCTGAAACACAAAAAGTAAAGCAAAAAAAGAGCGGTAA 3240
Db 3181 AAGCTGAACGACCTGCTGAAACACAAAAAGTAAAGCAAAAAAAGAGCGGTAA 3240
Qy 3241 AAAGGAGTGTCTGATCCCTGCTGATCAAGCCTGTAAAGCATTTGAAGCCAAAC 3300
Db 3241 AAAGGAGTGTCTGATCCCTGCTGATCAAGCCTGTAAAGCATTTGAAGCCAAAC 3300
Qy 3301 TTGAGGTATTTGATGCCCCACAGCAATCGAAAAAGATGCTAGCTAAAGAGAGCGG 3360
Db 3301 TTGAGGTATTTGATGCCCCACAGCAATCGAAAAAGATGCTAGCTAAAGAGAGCGG 3360
Qy 3361 AAAAACAAGCAAAACAAAAAGCTTGATCGAGCCGTATTAAGAGTGTATCAAGAT 3420
Db 3361 AAAAACAAGCAAAACAAAAAGCTTGATCGAGCCGTATTAAGAGTGTATCAAGAT 3420
Qy 3421 TATCTGACAGTAATAGTATGCTTCTGTTCAAGATGAATAGATGCTTTTGTAG 3480
Db 3421 TATCTGACAGTAATAGTATGCTTCTGTTCAAGATGAATAGATGCTTTTGTAG 3480
Qy 3481 ATCAAGCAATCTGCGCTGTGACAAATATGCGACAGATAAAGAGCGTATGATCTG 3540
Db 3481 ATCAAGCAATCTGCGCTGTGACAAATATGCGACAGATAAAGAGCGTATGATCTG 3540
Qy 3541 ATGCGTCCGCTGATTAAGACAGAAACGAATTAAGTGGGGTGGCAAAAG 3600
Db 3541 ATGCGTCCGCTGATTAAGACAGAAACGAATTAAGTGGGGTGGCAAAAG 3600
Qy 3601 CCTAGCTAATGAGCAATGAGGAGCTTTCTGCAATGCGTTCAAGATTAACCTTTG 3660
Db 3601 CCTAGCTAATGAGCAATGAGGAGCTTTCTGCAATGCGTTCAAGATTAACCTTTG 3660
Qy 3661 ATGAAACAGTTAAAAATCAAGCAATTAAGATGATGCTGCGTTTCCCATATTCAT 3720
Db 3661 ATGAAACAGTTAAAAATCAAGCAATTAAGATGATGCTGCGTTTCCCATATTCAT 3720
Qy 3721 GGGGCGATTAGCAATTTGGTGTAAACGTGGGAAAGCGGATCAAGCAATTAAGCTG 3780
Db 3721 GGGGCGATTAGCAATTTGGTGTAAACGTGGGAAAGCGGATCAAGCAATTAAGCTG 3780
Qy 3781 AAGAAACAAGCCGAAATTAATCAAGAAACGATTAATAGCGGTGAATGCAAGTATC 3840
Db 3781 AAGAAACAAGCCGAAATTAATCAAGAAACGATTAATAGCGGTGAATGCAAGTATC 3840
Qy 3841 AGTTCCGTTAGGGCAATTTGGCAATTAAGCTTATTTGGAGTTAATCGCTATTTATG 3900
Db 3841 AGTTCCGTTAGGGCAATTTGGCAATTAAGCTTATTTGGAGTTAATCGCTATTTATG 3900
Qy 3901 AACGTGAATTAATCAATCTGAGAGAGTGAAGTGAAGAGCGTACGCTTGAATTAATC 3960
Db 3901 AACGTGAATTAATCAATCTGAGAGAGTGAAGTGAAGAGCGTACGCTTGAATTAATC 3960
Qy 3961 GCTAATAATGCTGCAATGAGTGAATTAATTAATTAATCGACAGATTAATCAAGCTTA 4020
Db 3961 GCTAATAATGCTGCAATGAGTGAATTAATTAATTAATCGACAGATTAATCAAGCTTA 4020
Qy 4021 AGCTTAATTTCTGCAATTAATGATGATGTTCAAGCGCTAAGCAAGTGA 4080
Db 4021 AGCTTAATTTCTGCAATTAATGATGATGTTCAAGCGCTAAGCAAGTGA 4080
Qy 4081 ATCTGACGAGTGTGACACACATTTGAGGATTAATGAGAAAGAGGAGTGAAGG 4140
Db 4081 ATCTGACGAGTGTGACACACATTTGAGGATTAATGAGAAAGAGGAGTGAAGG 4140
Qy 4141 CAGAAATTTTACATTTCCAAATTTCCGCTTTATCTCAAAATCTCAAGTGTCACACTG 4200
Db 4141 CAGAAATTTTACATTTCCAAATTTCCGCTTTATCTCAAAATCTCAAGTGTCACACTG 4200
Qy 4201 GCAAAACGAAATGAGGAGTGAATTTGGGCTATGCTGTAATTAATCAATTAATTT 4260

```

```

Db 4201 GCAAAACGAAATGAGGAGTGAATTTGGGCTATGCTGTAATTAATCAATTAATTT 4260
Qy 4261 ATGCTTATTTGATTAAGCAAGTGGGTGAGATCAGATCCACCTTTCTTATTCATTAAT 4319
Db 4261 ATGCTTATTTGATTAAGCAAGTGGGTGAGATCAGATCCACCTTTTATTCATTAAT 4319

RESULT 2
US-10-645-655-1
Sequence 1, Application US/10645655
GENERAL INFORMATION:
APPLICANT: St. Geme III, Joseph W.
Falkow, Stanley
TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Flehr, Hohnbach, Test, Albrighton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/645,655
FILING DATE: 20-Aug-2003
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/296,791
FILING DATE: 25-Aug-1994
ATTORNEY/AGENT INFORMATION:
NAME: Treccatlin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4319 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
FEATURE:
NAME/KEY: CDS
LOCATION: 60..4241
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-645-655-1

Query Match 100.0%; Score 43.9; DB 53; Length 4319;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATATGCTTTAAGTATGATTTTAAATGCAAAAAATTAATTAATTAATCAATTA 60
Db 1 TCATATGCTTTAAGTATGATTTTAAATGCAAAAAATTAATTAATTAATTAATCAATTA 60
Qy 61 TGAATAAAAGCTATTTTGGCTTAATTTTAAACGCTGCAATTCATTAGGAGTAGTAT 120
Db 61 TGAATAAAAGCTATTTTGGCTTAATTTTAAACGCTGCAATTCATTAGGAGTAGTAT 120
Qy 121 CGAAGCGGTGCGTGTACACTTATTTGGGATTTGATTAACAAATATATGATTTTG 180
Db 121 CGAAGCGGTGCGTGTACACTTATTTGGGATTTGATTAACAAATATATGATTTTG 180
Qy 181 CGAAGAAATTAAGGAAATTTACAGTTTGGGCTCAAAATATTAAGTTATTAACAAACAG 240

```


Db	181	CCGAGATTAAGGGAAGTTCAAGTTGCGCTCAAAATATATTAAGTTATTAACAAACAG	240
Qy	241	GGCAATTAAGTGGCAATCAATGACAAAAGCCCGATGATTTTCTGAGTGCAC	300
Db	241	GGCAATTAAGTGGCAATCAATGACAAAAGCCCGATGATTTTCTGAGTGCAC	300
Qy	301	GTAAAGCGCTGGCAAGCTTGGTGAATTAATATATATGAGGCTGGACATACGTAG	360
Db	301	GTAAAGCGCTGGCAAGCTTGGTGAATTAATATATATGAGGCTGGACATACGTAG	360
Qy	361	GATATACAGATGTTGATTTTGTGACAGAGGAAACACCCGATCAATCTGTTACTT	420
Db	361	GATATACAGATGTTGATTTTGTGACAGAGGAAACACCCGATCAATCTGTTACTT	420
Qy	421	ATAAGATTGTAAGAAATTAATCAAAAAAGATTAATTAATCTCTTAAGAGACAT	480
Db	421	ATAAGATTGTAAGAAATTAATCAAAAAAGATTAATTAATCTCTTAAGAGACAT	480
Qy	481	ACCATATCCAGATTAATTAATCTTAAGAGAGCGCTCCAAATTAATGATGCTTGA	540
Db	481	ACCATATCCAGATTAATTAATCTTAAGAGAGCGCTCCAAATTAATGATGCTTGA	540
Qy	541	ATATGAATGCGAGTACTTATTCAGATTAACAAATATCCAGAACTGTTGATCGCT	600
Db	541	ATATGAATGCGAGTACTTATTCAGATTAACAAATATCCAGAACTGTTGATCGCT	600
Qy	601	CTGAGAGCGAGTTTGGGAAATGATCAAGACAAAGCGACAAAGTTCGCTGATATC	660
Db	601	CTGAGAGCGAGTTTGGGAAATGATCAAGACAAAGCGACAAAGTTCGCTGATATC	660
Qy	661	ATTATCTGACAGCTGGCAATACACACATCAGCTGGAGCAGTAAATGATTCGAT	720
Db	661	ATTATCTGACAGCTGGCAATACACACATCAGCTGGAGCAGTAAATGATTCGAT	720
Qy	721	TGGAGAGCGATGTTGTTAAACCGGAGAAATATGCTCCATTAACGATGAGCTCAAG	780
Db	721	TGGAGAGCGATGTTGTTAAACCGGAGAAATATGCTCCATTAACGATGAGCTCAAG	780
Qy	781	GGGACAGTGGTCTCCGATGTTTATTAATGATGAGAAACAAATGTTAATATG	840
Db	781	GGGACAGTGGTCTCCGATGTTTATTAATGATGAGAAACAAATGTTAATATG	840
Qy	841	GGATATTAAGGGAAGGCAACCTTTGAAAGGCAAAAGAAATGAGTTTCAATGCTGCA	900
Db	841	GGATATTAAGGGAAGGCAACCTTTGAAAGGCAAAAGAAATGAGTTTCAATGCTGCA	900
Qy	901	AATCTTATTTGATGAATTTTGGAAAGATTTTCAATCACTCACTTACACCGAGCTG	960
Db	901	AATCTTATTTGATGAATTTTGGAAAGATTTTCAATCACTCACTTACACCGAGCTG	960
Qy	961	GTAAATGAGTGAACAATTAAGTAAATGATTAATGTCAGGAGCTTATTAATCAAT	1020
Db	961	GTAAATGAGTGAACAATTAAGTAAATGATTAATGTCAGGAGCTTATTAATCAAT	1020
Qy	1021	CAGGAATACCATCAAGAAATTAATTAAGTGAACAAATTAATGATTTTCACTTGAAGGA	1080
Db	1021	CAGGAATACCATCAAGAAATTAATTAAGTGAACAAATTAATGATTTTCACTTGAAGGA	1080
Qy	1081	AGGATTAAGTGAATCTTAATGATTAAGTGAACAAATTAATGATTTTCACTTGAAGGA	1140
Db	1081	AGGATTAAGTGAATCTTAATGATTAAGTGAACAAATTAATGATTTTCACTTGAAGGA	1140
Qy	1141	ATGAGAAAGGCTTAATTTTAATGATTAAGTGAACAAATTAATGATTTTCACTTGAAGGA	1200
Db	1141	ATGAGAAAGGCTTAATTTTAATGATTAAGTGAACAAATTAATGATTTTCACTTGAAGGA	1200
Qy	1201	TAAACAAAGGAGCTGCTTATTTTGAAGGATTTTAAACAGATCTTCAAAATCTTA	1260
Db	1201	TAAACAAAGGAGCTGCTTATTTTGAAGGATTTTAAACAGATCTTCAAAATCTTA	1260
Qy	1261	ACCAAACTTGGCAAGAGCTGGCATACATGTAAGTGAAGAAATAGCAACGTTACTTGAAG	1320
Db	1261	ACCAAACTTGGCAAGAGCTGGCATACATGTAAGTGAAGAAATAGCAACGTTACTTGAAG	1320
Qy	1321	TAAATGCGTGAACATGATCGATTTTCTAAATTTGGTAAAGAAATGACGTTCAAG	1380
Db	1321	TAAATGCGTGAACATGATCGATTTTCTAAATTTGGTAAAGAAATGACGTTCAAG	1380
Qy	1381	CCAAAGGGAATTAAGGTTGATCAGCTGAGCGATGTAAGTCAATTTTGGAGCAGC	1440
Db	1381	CCAAAGGGAATTAAGGTTGATCAGCTGAGCGATGTAAGTCAATTTTGGAGCAGC	1440
Qy	1441	AGGCAAGATCAAGGCAACAAAGCCTTAATGATTAAGTGGTGAAGGAGGAGGAG	1500
Db	1441	AGGCAAGATCAAGGCAACAAAGCCTTAATGATTAAGTGGTGAAGGAGGAGGAG	1500
Qy	1501	GACCTTCAATTAAGATTAAGTAAACAAATTAAGTGAACGATTAATTTTGGCTTTC	1560
Db	1501	GACCTTCAATTAAGATTAAGTAAACAAATTAAGTGAACGATTAATTTTGGCTTTC	1560
Qy	1561	GAGCTTCAATTAAGATTAAGTAAACAAATTAAGTGAACGATTAATTTTGGCTTTC	1620
Db	1561	GAGCTTCAATTAAGATTAAGTAAACAAATTAAGTGAACGATTAATTTTGGCTTTC	1620
Qy	1621	AGGAGGAGGCAATGATGTAACATTAATTAACATCAAGCCTTAATGATTAAG	1680
Db	1621	AGGAGGAGGCAATGATGTAACATTAATTAACATCAAGCCTTAATGATTAAG	1680
Qy	1681	GGAACGAAGCATTTGTTCTACCTAATGAAATTAATTAATTAATTAATTAATTAAT	1740
Db	1681	GGAACGAAGCATTTGTTCTACCTAATGAAATTAATTAATTAATTAATTAATTAAT	1740
Qy	1741	AAATGCTCAACAGTGGTGGTGGGAAACAGTAAATTAATTAATTAATTAATTAAT	1800
Db	1741	AAATGCTCAACAGTGGTGGTGGGAAACAGTAAATTAATTAATTAATTAATTAAT	1800
Qy	1801	ACCTTATTAATTAACCAACAGAGATGCTATTTGCTACTTCAAGTGGTGAAT	1860
Db	1801	ACCTTATTAATTAACCAACAGAGATGCTATTTGCTACTTCAAGTGGTGAAT	1860
Qy	1861	TAAAGCGATTAATTAACCAACAGAGATGCTATTTGCTACTTCAAGTGGTGAAT	1920
Db	1861	TAAAGCGATTAATTAACCAACAGAGATGCTATTTGCTACTTCAAGTGGTGAAT	1920
Qy	1921	AGGCTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1980
Db	1921	AGGCTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1980
Qy	1981	TGCTGGGATTAACAGATTAACAGATTAACAGATTAACAGATTAACAGATTAACAG	2040
Db	1981	TGCTGGGATTAACAGATTAACAGATTAACAGATTAACAGATTAACAGATTAACAG	2040
Qy	2041	GCGGAAGTGGCTGGTCTGCAATGTTTCTCAATTAAGGGAATTAAGGATGAGCA	2100
Db	2041	GCGGAAGTGGCTGGTCTGCAATGTTTCTCAATTAAGGGAATTAAGGATGAGCA	2100
Qy	2101	ATATGCAATTAACAGATTAACAGATTAACAGATTAACAGATTAACAGATTAACAG	2160
Db	2101	ATATGCAATTAACAGATTAACAGATTAACAGATTAACAGATTAACAGATTAACAG	2160
Qy	2161	CAGATTAAGGATTAACAGATTAACAGATTAACAGATTAACAGATTAACAGATTAAC	2220
Db	2161	CAGATTAAGGATTAACAGATTAACAGATTAACAGATTAACAGATTAACAGATTAAC	2220
Qy	2221	ATGATTAAGGATTAACAGATTAACAGATTAACAGATTAACAGATTAACAGATTAAC	2280
Db	2221	ATGATTAAGGATTAACAGATTAACAGATTAACAGATTAACAGATTAACAGATTAAC	2280
Qy	2281	ATGATTAAGGATTAACAGATTAACAGATTAACAGATTAACAGATTAACAGATTAAC	2340
Db	2281	ATGATTAAGGATTAACAGATTAACAGATTAACAGATTAACAGATTAACAGATTAAC	2340
Qy	2341	CATTAAGCAACATGACCAACCAATTAAGCAATTAAGCAATTAAGCAATTAAGCA	2400
Db	2341	CATTAAGCAACATGACCAACCAATTAAGCAATTAAGCAATTAAGCAATTAAGCA	2400

2401 CGGTGATTAATGCAAACTTGAACGGTATATGCAATTAACGGATTCAGCTCAATTTCTT 2460
 2402 CGGTGATTAATGCAAACTTGAACGGTATATGCAATTAACGGATTCAGCTCAATTTCTT 2460
 2461 TAAAAACAGCCATTTTTCGACCAAAATTCAGGAGACAAAGGACAAACAGTGAAGTGG 2520
 2462 TAAAAACAGCCATTTTTCGACCAAAATTCAGGAGACAAAGGACAAACAGTGAAGTGG 2520
 2521 AAAATGCACTGGAACATGCGTACGCTACTATCATTAAGAAATTTAAAGCTTAATACA 2580
 2522 AAAATGCACTGGAACATGCGTACGCTACTATCATTAAGAAATTTAAAGCTTAATACA 2580
 2581 GTACGATCACTGTAATTAATTCAGCTTATTCAGTACTCAACATTAAGCCAGTCCGCTT 2640
 2582 GTACGATCACTGTAATTAATTCAGCTTATTCAGTACTCAACATTAAGCCAGTCCGCTT 2640
 2641 CATTAAGACGGAACCAACGCGTACATTCGAGACATTCGTTCAACATTCGAGTGA 2700
 2642 CATTAAGACGGAACCAACGCGTACATTCGAGACATTCGTTCAACATTCGAGTGA 2700
 2701 ATGTTAATTTGATGAGGCAAGGACATTCGATTTTCACTTATTTTGGCTAATAAA 2760
 2702 ATGTTAATTTGATGAGGCAAGGACATTCGATTTTCACTTATTTTGGCTAATAAA 2760
 2761 GCGATTAATTTAATTAATTCATCAATGACGCTGAGGAGATTAATTAATTCGTTGCA 2820
 2762 GCGATTAATTTAATTAATTCATCAATGACGCTGAGGAGATTAATTAATTCGTTGCA 2820
 2821 CAGGCAAAAGCCGGAACCCCTTGAGCAATTAATTTGTTGGAAGCAAAAGATTAATCA 2880
 2822 CAGGCAAAAGCCGGAACCCCTTGAGCAATTAATTTGTTGGAAGCAAAAGATTAATCA 2880
 2881 CTTTATGATTAAGTCAATTTTCACTTATTAAGAAATGACAGCTGATGAGGATTAAC 2940
 2882 CTTTATGATTAAGTCAATTTTCACTTATTAAGAAATGACAGCTGATGAGGATTAAC 2940
 2941 GTTATTAATTTAGTGAAGATTAATTCGCTTATTCGCTTATTAACCCATTAAGAGCAG 3000
 2942 GTTATTAATTTAGTGAAGATTAATTCGCTTATTCGCTTATTAACCCATTAAGAGCAG 3000
 3001 AATTCGCAATGATTTAGTGAAGATTAATTCGCTTATTCGCTTATTAACCCATTAAG 3060
 3002 AATTCGCAATGATTTAGTGAAGATTAATTCGCTTATTCGCTTATTAACCCATTAAG 3060
 3061 TTGAACGATGCTTAAACCAACAGGCTGAGCAAAAGTGGTCAAGAAAGAGCAGGA 3120
 3062 TTGAACGATGCTTAAACCAACAGGCTGAGCAAAAGTGGTCAAGAAAGAGCAGGA 3120
 3121 AGCAGGCTTCTGATACCTGCTGATCAAAAGCTGTTAAAGCATTAAGAGCAGAAC 3180
 3122 AGCAGGCTTCTGATACCTGCTGATCAAAAGCTGTTAAAGCATTAAGAGCAGAAC 3180
 3181 AAGCTGAATGATGCTGGAACCAACAAAGTGAAGCAAAAGTGAAGCAAAAGTGAAGCA 3240
 3182 AAGCTGAATGATGCTGGAACCAACAAAGTGAAGCAAAAGTGAAGCAAAAGTGAAGCA 3240
 3241 AAAAGAGCAGTCTTCTGATACCTGCTGATCAAAAGCTGTTAAAGCATTAAGAGCAG 3300
 3242 AAAAGAGCAGTCTTCTGATACCTGCTGATCAAAAGCTGTTAAAGCATTAAGAGCAG 3300
 3301 TTGAGGATTAATGATGCGGACCAACGATTCGGAAGAAAGTGTGCTCAAGAAAGAGG 3360
 3302 TTGAGGATTAATGATGCGGACCAACGATTCGGAAGAAAGTGTGCTCAAGAAAGAGG 3360
 3361 AAAAACAACGCAAAACAAAGCTTGAATGAGCCGTTATCAATTAAGTGAAGTGAAGT 3420
 3362 AAAAACAACGCAAAACAAAGCTTGAATGAGCCGTTATCAATTAAGTGAAGTGAAGT 3420
 3421 TATTCGCAAGTAAATGATGCTTCTGATCAAAAGCTGTTAAAGCATTAAGAGCAGT 3480
 3422 TATTCGCAAGTAAATGATGCTTCTGATCAAAAGCTGTTAAAGCATTAAGAGCAGT 3480
 3481 ATCAAGCAATTCCTGCGGTGAGCAAAATTCGACAGATTAAGAGCAGTATGATCTG 3540

3481 ATCAAGCAATTCCTGCGGTGAGCAAAATTCGACAGATTAAGAGCAGTATGATCTG 3540
 3541 ATGCTTCCGCTCTTATTCAGACCAAAAGCAATTAAGTGAAGTGAAGTGAAGTGAAG 3600
 3542 ATGCTTCCGCTCTTATTCAGACCAAAAGCAATTAAGTGAAGTGAAGTGAAGTGAAG 3600
 3601 CTTTATGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3660
 3602 CTTTATGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3660
 3661 ATGAACGATTAATTAATTCAGGACCAATTAAGATGATGAGGCTTATTCGCAATTAAT 3720
 3662 ATGAACGATTAATTAATTCAGGACCAATTAAGATGATGAGGCTTATTCGCAATTAAT 3720
 3721 GGGGCAATTAATTAATTCAGGACCAATTAAGATGATGAGGCTTATTCGCAATTAAT 3780
 3722 GGGGCAATTAATTAATTCAGGACCAATTAAGATGATGAGGCTTATTCGCAATTAAT 3780
 3781 AAGAACGATTAATTAATTCAGGACCAATTAAGATGATGAGGCTTATTCGCAATTAAT 3840
 3782 AAGAACGATTAATTAATTCAGGACCAATTAAGATGATGAGGCTTATTCGCAATTAAT 3840
 3841 AGTTCGTTAAGGCAATTCGAGCAATTCAGGCTTATTCGCAATTAATTCGCAATTAAT 3900
 3842 AGTTCGTTAAGGCAATTCGAGCAATTCAGGCTTATTCGCAATTAATTCGCAATTAAT 3900
 3901 AAGTGAATTAATTAATTCAGGACCAATTAAGATGATGAGGCTTATTCGCAATTAAT 3960
 3902 AAGTGAATTAATTAATTCAGGACCAATTAAGATGATGAGGCTTATTCGCAATTAAT 3960
 3961 GCTTATGATTAAGTCAATTCGAGGACCAATTAAGATGATGAGGCTTATTCGCAATTAAT 4020
 3962 GCTTATGATTAAGTCAATTCGAGGACCAATTAAGATGATGAGGCTTATTCGCAATTAAT 4020
 4021 AGCCTTATTCGCTGATTAATTCGAGGACCAATTAAGATGATGAGGCTTATTCGCAAT 4080
 4022 AGCCTTATTCGCTGATTAATTCGAGGACCAATTAAGATGATGAGGCTTATTCGCAAT 4080
 4081 ATTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4140
 4082 ATTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4140
 4141 CAGAAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 4200
 4142 CAGAAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 4200
 4201 GCAAAAGCAAAATGAGGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 4260
 4202 GCAAAAGCAAAATGAGGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 4260
 4261 ATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4320
 4262 ATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4320

RESULT 3
 US-10-687-046-1
 Sequence 1, Application US/10687046
 GENERAL INFORMATION:
 APPLICANT: St. Geme, Joseph W.
 TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 FILE REFERENCE: A-59941-1/RT/DCF/DR
 CURRENT APPLICATION NUMBER: US/10/687,046
 CURRENT FILING DATE: 2003-10-15
 PRIOR APPLICATION NUMBER: US/10/080,505
 PRIOR FILING DATE: 2002-02-22
 PRIOR APPLICATION NUMBER: US 08/296,791
 PRIOR FILING DATE: 1994-10-25
 PRIOR APPLICATION NUMBER: US 09/839,996
 PRIOR FILING DATE: 2001-04-20
 NUMBER OF SEQ ID NOS: 58
 SOFTWARE: PatentIn version 3.1

QY	1981	TTGGTGGGATCAAGATGGATCAACCGTATCATTTAAGCTGAAACCTTCCAAATTAAG	204
Db	1981	TTGGTGGGATCAAGATGGATCAACCGTATCATTTAAGCTGAAACCTTCCAAATTAAG	204
QY	2041	GCGAAGTGGGTGGTTCCTCGCAATGTTTCTTCAATTAGAGGAATTGACAGTCAGCA	2100
Db	2041	GCGAAGTGGGTGGTTCCTCGCAATGTTTCTTCAATTAGAGGAATTGACAGTCAGCA	2100
QY	2101	ATTAATCAAAATGCCAATTTGTTGTGTGCCAATCAACAAATACATTTGACGGCTT	2160
Db	2101	ATTAATCAAAATGCCAATTTGTTGTGTGCCAATCAACAAATACATTTGACGGCTT	2160
QY	2161	CAGATTGACAAAGATTACGACTGTGCAAAAAGTGATTTACCGATCAAAAGTTATTA	2220
Db	2161	CAGATTGACAAAGATTACGACTGTGCAAAAAGTGATTTACCGATCAAAAGTTATTA	2220
QY	2221	ATTCTATACAAAACCAATCAATGAGCTCTTAATTATACGTATATGCAAGGCCA	2280
Db	2221	ATTCTATACAAAACCAATCAATGAGCTCTTAATTATACGTATATGCAAGGCCA	2280
QY	2281	ATGTTAAAGTTTAGCAAACTTAAATGGCAATGTCACTTTAACAATCAAGCAATTA	2340
Db	2281	ATGTTAAAGTTTAGCAAACTTAAATGGCAATGTCACTTTAACAATCAAGCAATTA	2340
QY	2341	CATTAGCAACATGCCACCCCAATAGGCATATTCGACTTTCGACAAATTCAGCA	2400
Db	2341	CATTAGCAACATGCCACCCCAATAGGCATATTCGACTTTCGACAAATTCAGCA	2400
QY	2401	CGGTGATATATGCAAACTTGAAACGGTATGTGATTTAGGATTCAGTCAATTTCTT	2460
Db	2401	CGGTGATATATGCAAACTTGAAACGGTATGTGATTTAGGATTCAGTCAATTTCTT	2460
QY	2461	TAAAAACAGCCATTTTTCGACCAAAATTCAGGAGACAAAGGCAACAAGTACGTTGG	2520
Db	2461	TAAAAACAGCCATTTTTCGACCAAAATTCAGGAGACAAAGGCAACAAGTACGTTGG	2520
QY	2521	AAATATCGACTTGGACACATGCTTACGCAATCTACATTCGAAATTTTAACGCTAATTA	2580
Db	2521	AAATATCGACTTGGACACATGCTTACGCAATCTACATTCGAAATTTTAACGCTAATTA	2580
QY	2581	GTACGATCAACGTAAATATGAGTTATCTACATTAAGCTCAAAACATATAGCAAGTGGCGTT	2640
Db	2581	GTACGATCAACGTAAATATGAGTTATCTACATTAAGCTCAAAACATATAGCAAGTGGCGTT	2640
QY	2641	CATTAGACAGGAACCAAGGCCAATCTGGCAGAAACATCGTTTCAACACA TTGACAGTTA	2700
Db	2641	CATTAGACAGGAACCAAGGCCAATCTGGCAGAAACATCGTTTCAACACA TTGACAGTTA	2700
QY	2701	ATGCTAAATATGATGGGCAAGCACTTCCAAATTTACTCACTTATTTATTTGGCTAATAA	2760
Db	2701	ATGCTAAATATGATGGGCAAGCACTTCCAAATTTACTCACTTATTTATTTGGCTAATAA	2760
QY	2761	GGCATTAATTTAAATTTATTCGATATGACGCTGAGGGGAGATTACATATATCTGTGGACA	2820
Db	2761	GGCATTAATTTAAATTTATTCGATATGACGCTGAGGGGAGATTACATATATCTGTGGACA	2820
QY	2821	CAGGCAAAAGAACCCGAAACCCCTTAGAGCA TTACTTGGTGAAGCAAAAGATATATCAC	2880
Db	2821	CAGGCAAAAGAACCCGAAACCCCTTAGAGCA TTACTTGGTGAAGCAAAAGATATATCAC	2880
QY	2881	CGTTATACGATTAAGCTCAATTTACTTTGAAATATGACACGTTGATGCAAGTGCATTCAC	2940
Db	2881	CGTTATACGATTAAGCTCAATTTACTTTGAAATATGACACGTTGATGCAAGTGCATTCAC	2940
QY	2941	GTTATTAATTTGTAAGATATGATGGCAATTCGCTTGCAATACCCATTTAAAGACACAG	3000
Db	2941	GTTATTAATTTGTAAGATATGATGGCAATTCGCTTGCAATACCCATTTAAAGACACAG	3000
QY	3001	AATTGACAAATGATTTATGTAAGACAGACAGACAGACAAATTAAAGACCCAAACAG	3060
Db	3001	AATTGACAAATGATTTATGTAAGACAGACAGACAGACAGACAAATTAAAGACCCAAACAG	3060
QY	3061	TTGAACCACTGCTTAAACACAAACAGTGTAGCCAAAAGTGGGCTTCAGAAAGACAGCA	3120

Db	3061	TTGAAACCAATGCTAAACAAACAAAGGTGAGCCAAAGAGCGGTCAAGAAAGACGCCA	3122
Qy	3121	GAAGACCGTTTCTGAAACCTGCTGATCAAAAGCTGTTAAACCATTAAGAACCCAAAC	3183
Db	3121	GAAGACCGTTTCTGAAACCTGCTGATCAAAAGCTGTTAAACCATTAAGAACCCAAAC	3180
Qy	3181	AAGCTGAATGATGCTGAAAACAAAAAAGTAAAGCAAAAACAAAAAAGTGGGTCA	3240
Db	3181	AAGCTGAATGATGCTGAAAACAAAAAAGTAAAGCAAAAACAAAAAAGTGGGTCA	3240
Qy	3241	AAGAGCAGTGTTCGTGATCCCTGCTTGATCAAAAGCCTGTTGCGATTGAAGACCGCAC	3300
Db	3241	AAGAGCAGTGTTCGTGATCCCTGCTTGATCAAAAGCCTGTTGCGATTGAAGACCGCAC	3300
Qy	3301	TTGAGGTTATGAAAGCCCAACAGCAATGGA AAAAGATGCTTACGTCAGAAAGACCGC	3360
Db	3301	TTGAGGTTATGAAAGCCCAACAGCAATGGA AAAAGATGCTTACGTCAGAAAGACCGC	3360
Qy	3361	AAAAACAACGCAACAAAAAAGCTTGATCAGCCGTTATCAATATGAGCGTTACAGAT	3420
Db	3361	AAAAACAACGCAACAAAAAAGCTTGATCAGCCGTTATCAATATGAGCGTTACAGAT	3420
Qy	3421	TATCTGCAACGTAAATAGTATGCTTCTGTCAGATGATTAAGATGCTTTTGTAG	3480
Db	3421	TATCTGCAACGTAAATAGTATGCTTCTGTCAGATGATTAAGATGCTTTTGTAG	3480
Qy	3481	ATCAGACCAATGCGCGGTGCAAAATATGCCACAGATTAAGAGCGTATGATTCG	3540
Db	3481	ATCAGACCAATGCGCGGTGCAAAATATGCCACAGATTAAGAGCGTATGATTCG	3540
Qy	3541	ATGCGTCCGTGCTTATCAGCAGCAAAAAAGCACTTACGTCAAATGAGGTTGCAAAAG	3600
Db	3541	ATGCGTCCGTGCTTATCAGCAGCAAAAAAGCACTTACGTCAAATGAGGTTGCAAAAG	3600
Qy	3601	CCTTACGTAATGACGAATGAGGCAAGTTTCTCGCAGACGCGTTCAGATTAATACCTTG	3660
Db	3601	CCTTACGTAATGACGAATGAGGCAAGTTTCTCGCAGACGCGTTCAGATTAATACCTTG	3660
Qy	3661	ATGAAACGGTTAAATAATACGCGAATTAAAGATATGTCGGGTTTGGCCCATATCAAT	3720
Db	3661	ATGAAACGGTTAAATAATACGCGAATTAAAGATATGTCGGGTTTGGCCCATATCAAT	3720
Qy	3721	GAGGCGAATTTACAAATTTGGTGTAAACGTGAGAACGGGATCAGTGCAGTAATAGCTG	3780
Db	3721	GAGGCGAATTTACAAATTTGGTGTAAACGTGAGAACGGGATCAGTGCAGTAATAGCTG	3780
Qy	3781	AAGAAACAAACCCAAAAATCATGAAAACCGATTAATATGAGGTGATGCAAGTTATC	3840
Db	3781	AAGAAACAAACCCAAAAATCATGAAAACCGATTAATATGAGGTGATGCAAGTTATC	3840
Qy	3841	AGTTCGGTTAAGGCAATGAGCAATGAGCTTATTTGAGAGTAAACGCTATTTATATG	3900
Db	3841	AGTTCGGTTAAGGCAATGAGCAATGAGCTTATTTGAGAGTAAACGCTATTTATATG	3900
Qy	3901	AAGGTGAATATATCATCTGAGAAAGTGAAGTGA AAAACCGCTAGCCTGCAATTAATC	3960
Db	3901	AAGGTGAATATATCATCTGAGAAAGTGAAGTGA AAAACCGCTAGCCTGCAATTAATC	3960
Qy	3961	GCTAATAAGCTGCAATGAGTATATATCACTTATCTCGACAGATTAATATACGCTTA	4020
Db	3961	GCTAATAAGCTGCAATGAGTATATATCACTTATCTCGACAGATTAATATACGCTTA	4020
Qy	4021	AACCTTATTTCTCGCAATTTATGATATTTTCAAGCGCTAAGTACAAACACGATTA	4080
Db	4021	AACCTTATTTCTCGCAATTTATGATATTTTCAAGCGCTAAGTACAAACACGATTA	4080
Qy	4081	ATCTCAGGTGTGCAACACCAATTTGACGTTATGCAAAAAGAGTGGATTAAAG	4140
Db	4081	ATCTCAGGTGTGCAACACCAATTTGACGTTATGCAAAAAGAGTGGATTAAAG	4140
Qy	4141	CAGAAATTTACATTCGAATTTCCGCTTTATCTCAAAATCTCAAGGTTCAACAATCG	4200
Db	4141	CAGAAATTTACATTCGAATTTCCGCTTTATCTCAAAATCTCAAGGTTCAACAATCG	4200

Db 4141 CAGAAATTTTACATTCCTCAATTTCCGCTTTATATCTCAAAATCTCAGGTTGCAACTCG 4200
 Qy 4201 GCAACAGCAAAATATGCGCGCTGTAATATGCGCTATCGTTGGTAAATCAACATATTTT 4260
 Db 4201 GCAACAGCAAAATATGCGCGCTGTAATATGCGCTATCGTTGGTAAATCAACATATTTT 4260
 Qy 4261 ATCGTTATTTGATTAACAGAGTGGGTGATGATGATCCACCTTTTATTCATATAT 4319
 Db 4261 ATCGTTATTTGATTAACAGAGTGGGTGATGATGATCCACCTTTTATTCATATAT 4319

RESULT 4

PCT-US03-17092-520/C
 ; Sequence 520, Application PC/TUS0317092
 ; GENERAL INFORMATION:
 ; APPLICANT: BAKLETZ et al.
 ; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF H. INFLUENZAE
 ; FILE REFERENCE: 28335/38815
 ; CURRENT APPLICATION NUMBER: PCT/US03/17092
 ; NUMBER OF SEQ ID NOS: 588
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 520
 ; LENGTH: 8439
 ; TYPE: DNA
 ; ORGANISM: H. influenzae
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (8305)..(8309)
 ; OTHER INFORMATION: n = a, c, g, or t
 ; NAME/KEY: misc_feature
 ; LOCATION: (8427)..(8427)
 ; OTHER INFORMATION: n = a, c, g, or t
 PCT-US03-17092-520

Query Match 73.9%; Score 3193.4; DB 1; Length 8439;
 Best Local Similarity 85.3%; Pred. No. 0;
 Matches 3710; Conservative 0; Mismatches 576; Indels 61; Gaps 11;

Qy 1 TCATAGTGTGTTAATCTA-GTATTTTAAACGAAATTTACTTAAATTAATTAACAT 59
 Db 6039 TCATAGTGTGTTAATCTA-GTATTTTAAACGAAATTTACTTAAATTAATTAACAT 5980
 Qy 60 ATGAAAAAACTGTATTTGTCTTAATTTTAAACGCTGCACTTCAATAGGATAGTA 119
 Db 5979 ATGAAAAAACTGTATTTGTCTTAATTTTAAACGCTGCACTTCAATAGGATAGTA 5920
 Qy 120 TCGCAAGCGTGGCTGCTGCACTATTTTGGGATGATGATACCAATATATCGATTTT 179
 Db 5919 TCGCAAGCGTGGCTGCTGCACTATTTTGGGATGATGATACCAATATATCGATTTT 5860
 Qy 180 GCCGAGATTAAGGAGAGTTCACAGTTGGGCTCAAAATTTAGGTTTATTAACAA 239
 Db 5859 GCTGAGATTAAGGAGAGTTCACAGTTGGGCTCAAAATTTAGGTTTATTAACAA 5800
 Qy 240 GGGCAATTAAGTGGCAATCAATGACAAAGCCCGATGATATATTTTCTGATGCTCA 299
 Db 5799 GGGCAATTAAGTGGCAATCAATGACAAAGCCCGATGATATATTTTCTGATGCTCA 5740
 Qy 300 CGTAAAGCGGTGGCGCTGCTGTAATTAATTAATTTGAGCGTGGCACTATAGCTA 359
 Db 5739 CGTAAAGCGGTGGCGCTGCTGTAATTAATTAATTTGAGCGTGGCACTATAGCTA 5680
 Qy 360 GATATATACAGATGTTGATTTTGTGAGAGGAGAAACCCGATCAACATGCTTTACT 419
 Db 5679 GATATATACAGATGTTGATTTTGTGAGAGGAGAAACCCGATCAACATGCTTTACT 5620
 Qy 420 TATAAGATTGTAACGAAATTAATCAAAAAATTAATTAATCAATCTTATAGAGAGAT 479
 Db 5619 TATCAATATGTTAAACGTAATTAATTAATCAAACTGATCGAAGATCATATATAGGAGAT 5560
 Qy 480 TACCATATCCAGATTAATCAATTAATTCGTTACAGAGGCGCTCAATGATATGACTTCG 539

Db 5559 TACCATATCCAGATTAATCAATTAATTCGTTACAGAGGCGCTCAATGATATGACTTCG 5500
 Qy 540 AATATGATGCAAGTACTTATCAATGAAACAAATATCCGAAACGCTTCTATCGCC 539
 Db 5499 AGTATGAAAGGACAGACTATGCAACAAAGAAATATCCGATCGCTGATTTGCT 5440
 Qy 600 TCTGACGCGAGTTTGGGCAATATGATCAAGCAAAAGGCAACAGTTGCGGTGATAT 659
 Db 5439 TCTGATGCAATTTTGGGCAATATGATCAAGCAAAAGGCAACAGTTGCGGTGATAT 5380
 Qy 660 CATATCTGACAGCTGGCAATATGATCAAGCAATGCGGTGAGCGTATATGATTTGAT 719
 Db 5379 CATATCTGACAGCTGGCAATATGATCAAGCAATGCGGTGAGCGTATATGATTTGAT 5320
 Qy 720 TTGGAGGCGAGTCTGTTAAAGGAGAAATGTCATTCGATTCGAGGCTGCAAG 779
 Db 5319 TTGGAGGCGAGTCTGTTAAAGGAGAAATGTCATTCGATTCGAGGCTGCAAG 5260
 Qy 780 GGGGACAGTGTCTCGATGTTTATTTATGATGCTGAAAAAATAATGTTAA 839
 Db 5259 GGGGACAGTGTCTCGATGTTTATTTATGATGCTGAAAAAATAATGTTAA 5200
 Qy 840 GGGATTTAGGGGAAAGCAACCTTTTGAAGCAAAAGAAATGGTTTCAATTTGCTCC 899
 Db 5199 GGGATTTAGGGGAAAGCAACCTTTTGAAGCAAAAGAAATGGTTTCAATTTGCTCC 5140
 Qy 900 AATCTATTTTATGTAATTTTTCGAAAGAGAT-TTACATCATGCTTATACCCGAGC 958
 Db 5139 AATCTATTTTATGTAATTTTTCGTAATTTTTCGTAATTTTTCGTAATTTTTCGTA 5080
 Qy 959 TGGTATGAGTGTACCAATTAATGTA-----ATGATATGTCAGGGGCTATAT 1010
 Db 5079 TGGTATGAGTGTACCAATTAATGTA-----ATGATATGTCAGGGGCTATAT 5020
 Qy 1011 ACTGCAATATGAGATATCATGCAAAATTAATTAATGCTTATGCAATATGATTTACT 1070
 Db 5019 ACTGCAATATGAGATATCATGCAAAATTAATTAATGCTTATGCAATATGATTTACT 4960
 Qy 1071 TGGAAAGAGAGTAAGTCTATATGCAATATGAGGAGCTTAATTTATTCGCA 1130
 Db 4959 TGGAAAGAGAGTAAGTCTATATGCAATATGAGGAGCTTAATTTATTCGCA 4900
 Qy 1131 CGTTTAAACATGAGAAACGCTATATTTATGATCAAAAAAAGATCATTAATCTTC 1190
 Db 4905 GATTAAGTACGCGGAAACCTTTATTTATGATTAAGAAAAAGGGGCTATACCTTT 4846
 Qy 1191 GATCTGACATTAACAAAGGCGGCTGCTTATTTTGAAGGTAATTTTACAGATCT 1250
 Db 4845 GATCTGACATTAACAAAGGCGGCTGCTTATTTTGAAGGTAATTTTATAGTTCTG 4786
 Qy 1251 CCAATTTCAACCAACTTGGGCAAGCTGCAATGATGATGAAATATGACACGCT 1310
 Db 4785 CCAATTTCAACCAACTTGGGCAAGCTGCAATGATGATGAAATATGACACGCT 4726
 Qy 1311 ACTTGAAGATTAATGCGTGGCAATGATGATGATGATGATGATGATGATGATGATG 1370
 Db 4725 ACTTGAAGATTAATGCGTGGCAATGATGATGATGATGATGATGATGATGATGATG 4666
 Qy 1371 CAGCTTCAAGCAAAAGGCAATTAAGGTTGATGAGGCTGATGATGATGATGATGAT 1430
 Db 4665 CAGCTTCAAGCAAAAGGCAATTAAGGTTGATGAGGCTGATGATGATGATGATGATG 4606
 Qy 1431 TTGGAGGAGAGGAGAGTCAAGCAACCAAGCCCTTATGATGATGATGATGATGATG 1490
 Db 4605 TTGGAGGAGAGGAGAGTCAAGCAACCAAGCCCTTATGATGATGATGATGATGATG 4546
 Qy 1491 AGGAGAGAGGAGAGTCAATTAAGATGATTAACCAATTTGATGATGATGATGATGAT 1550
 Db 4545 AGGAGAGAGGAGAGTCAATTAAGATGATTAACCAATTTGATGATGATGATGATGAT 4486
 Qy 1551 TTGGGCTTCTGCTGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATG 1610

Db 4485 TTGGCTTTCGCTGCTGCTTGAATCTTAATGAGCAATTAATTAACCTTTAAAGATAC 4426
 Qy 1611 CAATAACGACGAGGAGGCAATGATGTTGACCATTAATACCTCAAGCCGTAAATGC 1670
 Db 4425 CAATAACGACGAGGAGGCAATGATGTTGACCATTAATACCTCAAGCCGTAAATGC 4366
 Qy 1671 ACTATTAAGTGAAGCAAGATGTTCTACCTTAATGAAATTAATTAATAAATCTAT 1730
 Db 4365 ACTATTAAGTGAAGCAAGATGTTCTACCTTAATGAAATTAATTAATAAATCTAT 4306
 Qy 1731 TACGAAAGAAATTCCTCAACGCTGTTGTTGCGAAACAGATTAATAATTAACAAAT 1790
 Db 4305 TACGAAAGAAATTCCTCAACGCTGTTGTTGCGAAACAGATTAATAATTAACAAAT 4246
 Qy 1791 GAGGATTAATTAACCTTAATTAATTAACCAACAGAAAGATGCTTGTGCTACCTTCACT 1850
 Db 4245 GAGGATTAATTAACCTTAATTAATTAACCAACAGAAAGATGCTTGTGCTACCTTCACT 4186
 Qy 1851 GGTACAAATTAATAAGGCGATTAATCCCAACAAAGGTAATCTATTTTCAGCGGTAGA 1910
 Db 4185 GGTACAAATTAATAAGGCGATTAATCTCAAGAAAGGCGCATTTAGTGTGTTAGTGTGCG 4126
 Qy 1911 CCGACACCGACGCTCAACATCTTAATTAATAAAGCTGTTGCGAAATGGAAGTATACCA 1970
 Db 4125 CCGACACCGACGCTCAACATCTTAATTAATAAAGCTGTTGCGAAATGGAAGTATACCA 4072
 Qy 1971 CAAGCGAAATGCTGTTGCGATCAAGATGTTGCAACGCTACCTTAATTAAGCTGAAATCTC 2030
 Db 4071 CAAGCGAAATGCTGTTGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 4012
 Qy 2031 CAATTAAGGCGAAATGCTGTTGCGATGATGATGATGATGATGATGATGATGATGATGATG 2090
 Db 4011 CAATTAAGGCGAAATGCTGTTGCGATGATGATGATGATGATGATGATGATGATGATGATG 3952
 Qy 2091 AAGTCAAGATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 2150
 Db 3951 AAGTCAAGATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 3892
 Qy 2151 TGCAGCGCTTCAAGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2210
 Db 3891 TGCAGCGCTTCAAGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3832
 Qy 2211 AAAGTTTATTAATCTTAACCAAAACCAATTAATGCTCTTAATTAATTAATTAATTAAT 2270
 Db 3831 AAAGTTTATTAATCTTAACCAAAACCAATTAATGCTCTTAATTAATTAATTAATTAAT 3772
 Qy 2271 GCAAGCGCAATGTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAG 2330
 Db 3771 GCAAGCGCAATGTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAG 3712
 Qy 2331 AGCAATTTAATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 2390
 Db 3711 AGCAATTTAATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 3652
 Qy 2391 TCAACTGCAACGCTGATTAATGCAAACTTGAACGCTATGCTATGCTATGCTATGCTATGCT 2450
 Db 3651 TCAACTGCAACGCTGATTAATGCAAACTTGAACGCTATGCTATGCTATGCTATGCTATGCT 3592
 Qy 2451 CAATTTCTTTAAATAAAGCAATTTTGAACCAATTTGAAGGAGAGCAAAAGGAGCAAGCA 2510
 Db 3591 CAATTTCTTTAAATAAAGCAATTTTGAACCAATTTGAAGGAGAGCAAAAGGAGCAAGCA 3532
 Qy 2511 GTGACGTTGAAATGCACTTGAACCAATGCTTGAACCAATGCTTGAACCAATGCTTGAACCA 2570
 Db 3531 GTGACGTTGAAATGCACTTGAACCAATGCTTGAACCAATGCTTGAACCAATGCTTGAACCA 3472
 Qy 2571 CTAAATTAAGATGCAATGCTTGAACCAATGCTTGAACCAATGCTTGAACCAATGCTTGAAC 2630
 Db 3471 CTAAATTAAGATGCAATGCTTGAACCAATGCTTGAACCAATGCTTGAACCAATGCTTGAAC 3412
 Qy 2631 -----CGTCCGCTTCAATTAAGAGAGCAAAACCAAGCAATGCGAGCAATGCTTCC 2684
 Db 3411 CGCTGCTGCTGCTTCAATTAAGAGAGCAAAACCAAGCAATGCGAGCAATGCTTCC 3352

Qy 2685 AACACATGACAGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 2744
 Db 3351 AACACATGACAGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 3292
 Qy 2745 TTAATTTGCTTAAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2804
 Db 3291 TTAATTTGCTTAAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3232
 Qy 2805 TTAATTTGCTTAAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2864
 Db 3231 TTAATTTGCTTAAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3172
 Qy 2865 AGCAAGATTAATTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2924
 Db 3171 AGCAAGATTAATTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3112
 Qy 2925 GATGACGCTGATTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2984
 Db 3111 GATGACGCTGATTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3052
 Qy 2985 CCAATTAATTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3044
 Db 3051 CCAATTAATTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2992
 Qy 3045 TTAAGAGCAAAAGGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACG 3104
 Db 2991 TTAAGAGCAAAAGGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACG 2932
 Qy 3105 TTAAGAGCAAAAGGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACG 3164
 Db 2931 TTAAGAGCAAAAGGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACG 2881
 Qy 3165 GCATTAAGAGCAAAAGGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 3224
 Db 2880 GCATTAAGAGCAAAAGGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 2821
 Qy 3225 AAAAAGGCTGCTCAAAAGAGCA-----GTGTTTCTGATCCCTGCTGTTGAAC 3275
 Db 2820 AAAAAGGCTGCTCAAAAGAGCA-----GTGTTTCTGATCCCTGCTGTTGAAC 2761
 Qy 3275 AGCTGTTT-----CGATTAAGAGCAAGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGA 3332
 Db 2761 AGCTGTTT-----CGATTAAGAGCAAGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGA 2701
 Qy 3333 AAAGATGCTTGAATCAAGAGCAAGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 3392
 Db 2701 AAAGATGCTTGAATCAAGAGCAAGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 2650
 Qy 3393 CGTATTAATTAATGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 3452
 Db 2649 CGTATTAATTAATGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 2590
 Qy 3453 CAAGATTAATTAATGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 3512
 Db 2589 CAAGATTAATTAATGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 2530
 Qy 3513 GCAAGATTAATTAATGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 3572
 Db 2529 GCAAGATTAATTAATGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 2473
 Qy 3573 AACTTACGCTAAATGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 3632
 Db 2472 AACTTACGCTAAATGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 2413
 Qy 3633 TGCATTAATTAATGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 3692
 Db 2412 TGCATTAATTAATGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 2353
 Qy 3693 AAGATGCTGCTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 3752
 Db 2352 AAGATGCTGCTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAACGCTGTTGAAC 2293

QY 1 TCATATGTCGTTAACTA-GTATTTTATATCGAAAAATTACTTAATTAATTAACATT 59
Dp 6039 TCATATGTCGTTAACACGATATTTTATATCGAAAAATTACTTAATTAATTAACATT 598

Qy	69	ATGAAAAAACTGTAATTCGCTCTTAACTTTTAAACCGGTGATTTATATAGAGATGA	119
Db	5979	ATGAAAAAACTGTAATTCGCTCTTAACTTTTAAACCGGTGATTTATATAGAGATGA	592
Qy	120	TTCGAACGCTGGCTGGTCACTATTTTGGGATTTGATTACATATTAATCGATTTT	179
Db	5919	TTCGAACGCTGGCTGGTCACTATTTTGGGATTTGATTACATATTAATCGATTTT	586
Qy	180	GCGAGAAATTAAGGAAAGTTCACACTGCGGCTCAAAATATTAAGTTTATACAAACA	239
Db	5859	GCTGAGATTAAGGAAAGTTCACACTGCGGCTCAAAATATTAAGTTTATACAAACA	580
Qy	240	GGGCAATAGTTGCGACATCAATGACAAAGCCCGATGATTGATTTTCTGTAGTCA	299
Db	5799	GGGCAATAGTTGCGCAATCAATGACAAAGCCCGATGATTGATTTTCTGTAGTCA	5744
Qy	300	CGTAAAGCGGTGGCAGCTTGGTTAAATCAATATTTTGGAGCTGCGACATACGTA	359
Db	5739	CGTAAAGCGGTGGCAGCTTGGTTAAATCAATATTTTGGAGCTGCGACATACGTA	588
Qy	360	GGATATACAGATGTTGATTTGGTGCAGAGGAAAACAACCCGATCAATCGTTTAC	419
Db	5679	GGATATACATATGCTGATTTGGTGCAGAGGAAAACAACCCGATCAATCGTTTAC	562
Qy	420	TATTAAGTTTAAAGCAATTACTACAAAAGATATTTACATCTTATAGAGACAT	479
Db	5619	TATCAATATGTTAAACGTAATATTAATTAACCTGATCAGAGATCATATATAGGAGAT	556
Qy	480	TACCATATCCAGATTAACCTAAATTCGTTACAGAGGCGCCCAATGATATGACTGG	539
Db	5559	TACCATATCCCTCGTTACATTAATTTGTTACGAAAGTTCTCTTTGATATACGGA	550
Qy	540	AATATGATGAGTACATCTTATTCAGATAGAACAAATATCCGACAGTGTGATCGC	599
Db	5499	AGATATGAGGACGACATATGACAGACAAAGAAAATACCGAGGTGTGCGATTTGT	544
Qy	600	TTCGACAGCGACTTTGCGGAATGATCAAGACAAAGCCGACGAACTTCGCTGAT	659
Db	5439	TCGCGATCAATTTGCGGAATGACAAAGATTAAGGACCTGATACCTGAAAGCTTAT	538
Qy	660	CATTATCTGACAGCTGGCAATACACAAATCAGCGTGGAGAGGTATGATTTGAT	719
Db	5379	GATATCTTATTTGCGGGAATACCAATGCAACTGTGTGCGAAATGAGATTTAGT	532
Qy	720	TTCGAGAGCGAGTTGTGTAAGCGGAGAAATGTCTCATTAACGATTCAGGCTCAAG	779
Db	5319	TTTAGTGCGAGTGTGTAAGGAAAGATTTGACCTATGCGCAGCGAGATCATTT	526
Qy	780	GGGGAAGGTGTTCGCAATGTTATTAATGATGCTGAAAACAAAAATGTTAATAT	839
Db	5259	GCGAGAGGTGTTCCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT	5230
Qy	840	GGGATTTAAGGAGGCAACCTTTTGAAGGCAAGAAATGGTTTCAATTTGTTGCG	899
Db	5139	GGTATGCTGCTAGTGAACCAATCTTATAGTAAAGGAATCTTTACATAGTCTGT	514
Qy	900	AAATCTTATTTGATGAATTTTGGAAAGAT-TTACATATCACTTACACCCGAC	958
Db	5139	AAAAATATTTTGAATAATTTTCTCTCAATCTTAAACGCAATTTTGAATTCACA	5080
Qy	959	TGGTATAGATGATGACAACTTATGGA-----ATGATATGCTCAGCGGTCTATA	1010
Db	5079	ACTCTTAAATATCAAAATATGACATCTTCAATTTCAAAAATGATGGGTACGGC	5020
Qy	1011	ACTCGAAATCAGGAATACATCGAAATTAATTAATTCGTTAGCAATATGATTTACT	1070
Db	5019	ACTATTAATCTGACCTGTTAAATTTTAAAGTCAAACTGCTATTAATTAATTAATCT	4960
Qy	1071	TTTAAAGAGAGATTAAGTCTATTAATCTTATTAAGAGACTTATATTTATTTGCA	1130
Db	4959	ATTAAGGACGAGCTCTCTATCT-----ATGATGTTAAACAAATTTCTCC	4906

QY 1131 CGTTAAACATGAGAAAGCTATATTATGATCAAAACAGGATCATTAATCTC 1190
 DB 4905 GCATTAAGTCAAGCGAAAGCTTTATTTATGATTAAGAAAGGGGTGTAATCTT 4846
 QY 1191 GCATCTGACATTAACCAAGGGGGGGGTCTTATTTGAGGGTAAATTTACATATCT 1253
 DB 4845 AGTAAATATATCATCAAGGTGAGGGGGTGTATTTGAGGAGAAATTTATGTTGCG 4786
 QY 1251 CCAATTTCTTAACAACTTGCAAGAGCTGGGATACATGTAAGTGAATAAGACACGCT 1310
 DB 4785 CCAACCATATGAAAGCGAAAGGGGGGGAATATATCAGTGAAGAACACCTTA 4726
 QY 1311 ACTTGGAAGTAATGGCTGAGACATGATGCACTTTCTAAATTTGGTAAGGAACATG 1370
 DB 4725 ATTTGGAAGTCAATGGCTGAGAAATGATGCGCTTTCAAAATTTGGTAAGGAACATG 4666
 QY 1371 CAGCTTCAAGCAAAAGGGAATTAAGGTGATGATGAGTGAAGCGATGTTAAATGCT 1430
 DB 4665 CATATTCAAGCAAAAGGTAATAAATCTTAGGCTCAATGAGTGGCGATGTAAGTATC 4606
 QY 1431 TTGAGACAGAGGACAGACATCAAGGCAACAAAGCTTTAGTGAATGGCTTGGT 1490
 DB 4605 TTGAGACACAGAGGACAGATCAATTAATAAAGAGCTTTAGTGAATGGATTTGGT 4546
 QY 1491 AGCGGAGAGGAGCTGTTCAATTAAGATGATTAACATTTGATACGATTAATTTAT 1550
 DB 4545 AGCGGAGAGGAGCTGTTCAATTAAGATGATTAACATTTGATACGATTAATTTAT 4486
 QY 1551 TTGGCTTGT 1610
 DB 4485 TTGGCTTGT 4426
 QY 1611 CAAATTAAGGACGAGGGGCAATGATTTGAAACATTAATCAATCAATCAATCAAT 1670
 DB 4425 CAAATTAAGGACGAGGGGCAATGATTTGAAACATTAATCAATCAATCAATCAAT 4366
 QY 1671 ACTTATCTGGGAGAGAAAGCATTTGTTCTTACCTTAATGAAATTAATTAATTAAT 1730
 DB 4365 ACTTATCTGGGAGAGAAAGCATTTGTTCTTACCTTAATGAAATTAATTAATTAAT 4306
 QY 1731 TACAGAAAGAAATTTGCTCAATCAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1790
 DB 4305 TACAGAAAGAAATTTGCTCAATCAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4246
 QY 1791 GGGGATTAACCTTATTTATTAACCAACCAACCAACCAACCAACCAACCAACCAAC 1850
 DB 4245 GAGGATTAACCTTATTTATTAACCAACCAACCAACCAACCAACCAACCAACCAAC 4186
 QY 1851 GGTACAAATTTAAAGGAGATTAACCAACCAACCAACCAACCAACCAACCAACCAAC 1910
 DB 4185 GGTACAAATTTAAAGGAGATTAACCAACCAACCAACCAACCAACCAACCAACCAAC 4126
 QY 1911 CCGACACGACGCTCAATCAATCAATTAATTAAGGTGTGTGTGTGTGTGTGTGTGTGT 1970
 DB 4125 CCGACACGACGCTCAATCAATTAATTAAGGTGTGTGTGTGTGTGTGTGTGTGTGT 4072
 QY 1971 CAAGGCAATTTGT 2030
 DB 4071 CAAGGCAATTTGT 4012
 QY 2031 CAAATTAAGGAGAGAGGT 2090
 DB 4011 CAAATTAAGGAGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3952
 QY 2091 ACAGTCAAGCAATTAAGCAATTTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2150
 DB 3951 ACAGTCAAGCAATTAAGCAATTTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3892
 QY 2151 TGAAGCGCTTCAATGAGACAGATTAACCACTTGTCAAAAGTGAATTTAACGATACA 2210
 DB 3891 TGAAGCGCTTCAATGAGACAGATTAACCACTTGTCAAAAGTGAATTTAACGATACA 3832
 QY 2211 AAGTATTAATTTCTAATCAAAACCAATTAAGGTGTGTGTGTGTGTGTGTGTGTGT 2270

DB 3891 AAGTATTAATTTCTAATCAAAACCAATTAAGGTGTGTGTGTGTGTGTGTGTGTGT 3772
 QY 2271 GCAAGCGCAATTTGTAAGGTTTGAAGAACTTAATGCAATGTCACTTTAACCAATCAC 2330
 DB 3771 GCAAGCGCAATTTGTAAGGTTTGAAGAACTTAATGCAATGTCACTTTAACCAATCAC 3712
 QY 2331 AGCGATTTACATTTAGCAACATTTGCAACCAATTAAGCAATTTGCACTTTCCGACAT 2390
 DB 3711 AGCGATTTACATTTAGCAACATTTGCAACCAATTAAGCAATTTGCACTTTCCGACAT 3652
 QY 2391 TCAACTGCAACGAGTGAATTTGCAACCTTGAACGTTATGCACTTTAAGGATTTACGCT 2450
 DB 3651 TCAACTGCAACGAGTGAATTTGCAACCTTGAACGTTATGCACTTTAAGGATTTACGCT 3592
 QY 2451 CAATTTCTTTAAAAACCAAGCTTTTGTGCAACCAATTTGAGGAGACAAAGGACAAACA 2510
 DB 3591 CAATTTCTTTAAAAACCAAGCTTTTGTGCAACCAATTTGAGGAGACAAAGGACAAACA 3532
 QY 2511 GTGAGGTGGAATTTGGGACTTTGACATGCTTACGATTTCACTTGCAGATTTAAG 2570
 DB 3531 GTGAGGTGGAATTTGGGACTTTGACATGCTTACGATTTCACTTGCAGATTTAAG 3472
 QY 2571 CTAATTAACAGTACATCACTTTAATTTCACTTATTTCACTTATTTCACTTATTTCACT 2630
 DB 3471 CTAATTAACAGTACATCACTTTAATTTCACTTATTTCACTTATTTCACTTATTTCACT 3412
 QY 2631 -----CGTCCCGTGTATTAAGACGAAACCAAGCAACATTTGAGCAACATTTGCTTTC 2684
 DB 3411 CGTCCCGTGTATTTAAGACGAAACCAAGCAACATTTGAGCAACATTTGCTTTC 3352
 QY 2685 AACACATTAAGCAATTTGTAATTTGAGGAGGACATTTCACTTATTTCACTTATTTCACT 2744
 DB 3351 AACACATTAAGCAATTTGTAATTTGAGGAGGACATTTCACTTATTTCACTTATTTCACT 3292
 QY 2745 TATTTGCTTAAAGGCAATTAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 2804
 DB 3291 TATTTGCTTAAAGGCAATTAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 3232
 QY 2805 TATCTGTGTGCAACAGGACGAAACCGGAAACCTTTGAGCAATTTAATTTGCTTGA 2864
 DB 3231 TATCTGTGTGCAACAGGACGAAACCGGAAACCTTTGAGCAATTTAATTTGCTTGA 3172
 QY 2865 AGCAAGATTAATCAACGTTATCACTTACGTAATTTAATTTAATTTAATTTAATTTAATTT 2924
 DB 3171 AGCAAGATTAATCAACGTTATCACTTACGTAATTTAATTTAATTTAATTTAATTTAATTT 3112
 QY 2925 GATGAGGTGATTAAGTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 2984
 DB 3111 GATGAGGTGATTAAGTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 3052
 QY 2985 CCAATTAAGGAGCAATTTGCAATTTGTAAGGACGAAACGAAACGAAACGAAACGAAAC 3044
 DB 3051 CCAATTAAGGAGCAATTTGCAATTTGTAAGGACGAAACGAAACGAAACGAAACGAAAC 2992
 QY 3045 TTAGAAGCAAAAGATTTGAAACGCTGTTAAACCAAAAGGAGGAGGAGGAGGAGGAGG 3104
 DB 2991 TTAGAAGCAAAAGATTTGAAACGCTGTTAAACCAAAAGGAGGAGGAGGAGGAGGAGG 2932
 QY 3105 TGAAG 3164
 DB 2931 TGAAG 2881
 QY 3165 GATTTGAAGGCAAAAGGAG 3224
 DB 2881 GATTTGAAGGCAAAAGGAG 2821
 QY 3225 AAAAAAGTGGTCAAAAAAGCA-----GTGTTTCAATTTCCCGTGTGATGCA 3275
 DB 2821 AAAAAAGTGGTCAAAAAAGCA-----GTGTTTCAATTTCCCGTGTGATGCA 2761
 QY 3275 AGCTGTTT-----GCATTTGAAGGCACTTGAAGTGTGAGTGTGAGGAGAGAGAGAGAG 3332

```

Db 2760 AGCCAGTTAGAGCTATTAAAGCCGCACTTGAGTTATTAAATGCCCAACCGCAAGTGGCA 2701
Qy 3333 AAAATGTTCTAGCTCAAGAGAAAGCGAAACCAACGCAAAACAAAGATTGATCAGC 3392
Db 2760 AAAAGAGCTCAGCTCAA-----GAGAAACAAAGCAACAAAGAAAGATTGATCAGC 2850
Qy 3393 CGTATTCAAATAGTCCCTATTCAGAAATATCTGCAACAGTAATATGATGCTTCTGTT 3452
Db 2649 CGTATTCAAATAGTCCCTATTCAGAAATATCTGCAACAGTAATATGATGCTTCTGTT 2590
Qy 3453 CAAGATGATATGATGCTTCTTGTGATGATCAAGCAATCTGCGGTGAGCAAAATATC 3512
Db 2589 CAAGATGATATGATGCTTCTTGTGATGATCAAGCAATCTGCGGTGAGCAAAATATC 2530
Qy 3513 GCACAGATPAAAGACGCTATGATTCGATTCGATTCGCTTATGACAGAGAAACG 3572
Db 2529 GCACAGATPAAAGACGCTATGATTCGATTCGATTCGCTTATGACAGAGAAACG 2473
Qy 3573 AACTGACGTCAAATGGGGTGCAGAAAGCCTTAGCTAATGACAAATGGGGCACTTTTC 3632
Db 2472 AACTGACGTCAAATGGGGTGCAGAAAGCCTTAGCTAATGACAAATGGGGCACTTTTC 2413
Qy 3633 TCCGATAGCCGTTGAGATTAATACCTTGTATGACAGGTTAAATCAGCGCACTTAAAG 3692
Db 2412 TCCGATAGCCGTTGAGATTAATACCTTGTATGACAGGTTAAATCAGCGCACTTAAAG 2353
Qy 3693 ATGATGTCGGGTTTCCCAATATCAATGGGGCAATTTACATTTGGTGTAAACGTGGCA 3752
Db 2352 ATGATGTCGGGTTTCCCAATATCAATGGGGCAATTTACATTTGGTGTAAACGTGGCA 3752
Qy 3753 AGGGATTCAGTGGAGTAAATGGCTGAAAGCAAGCCGAAATTCATCGAAAGCG 3812
Db 2292 GCGGGAATTTAGTGAATPAAATGGCTGAAAGCAAGCCGAAATTCATCGAAAGCG 2233
Qy 3813 ATAAATTAATGCGTAATGCAAGTATCATGCTGCTTGGGCAATGGGCACTTCAGCT 3872
Db 2232 ATAAATTAATGCGTAATGCAAGTATCATGCTGCTTGGGCAATGGGCACTTCAGCT 2173
Qy 3873 TATTTGAGATTAATGCTATTTATGTAAGCTGAAATTAATCAATCGAGAAAGTGA 3932
Db 2172 TATTTGAGATTAATGCTATTTATGTAAGCTGAAATTAATCAATCGAGAAAGTGA 2113
Qy 3933 GTGAAAGCGCTACCGTTCGATTAATGCTATTAATGCTGATTCGATTAATGCA 3992
Db 2112 GTGAAAGCGCTACCGTTCGATTAATGCTATTAATGCTGATTCGATTAATGCA 2053
Qy 3993 TTTACTCCGACAGATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTGAT 4052
Db 2052 TTTACTCCGACAGATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTGAT 1993
Qy 4053 TCAGAGCTTAAGTAAACAGAGTAAATCTGACGCTGTCAGCAACCAATTTGAGCT 4112
Db 1992 TCAGAGCTTAAGTAAACAGAGTAAATCTGACGCTGTCAGCAACCAATTTGAGCT 1933
Qy 4113 TATTTGAGAAAGAGTGGATTTAAAGCGAAATTTTCAATTCGAAATTTCCGCTTT 4172
Db 1932 TATTTGAGAAAGAGTGGATTTAAAGCGAAATTTTCAATTCGAAATTTCCGCTTT 1873
Qy 4173 ATCTCAAAATCTCAAGGTTCAACAACGCGAAACAGCAAAATGGGGCGTGAATTTGCGC 4232
Db 1872 ATCTCAAAATCTCAAGGTTCAACAACGCGAAACAGCAAAATGGGGCGTGAATTTGCGC 1813
Qy 4233 TATCTGTTGTAATCAACATTAATTTATGTTATGATTAACAGAGTGGGTGAGATC 4292
Db 1812 TATCTGTTGTAATCAACATTAATTTATGTTATGATTAACAGAGTGGGTGAGATC 1753
Qy 4293 AGATCCGACCTTTTATTCGAATAT 4319
Db 1752 AAATCCGACCTTTTATTCGAATAT 1726

```

```

; Sequence 8, Application PC/TUS0105226
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/REF/DCE/DBR
; CURRENT APPLICATION NUMBER: PCT/US03/05226
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent version 3.1
; SEQ ID NO: 8
; LENGTH: 4305
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1702)..(1702)
; OTHER INFORMATION: "n" at position 1702 can be any base.
; NAME/KEY: CDS
; LOCATION: (1)..(4305)
; OTHER INFORMATION:
PCT-US03-05226-8

```

```

Query Match 73.6%; Score 3180.8; DB 1; Length 4305;
Best Local Similarity 85.7%; Pred. No. 0;
Matches 3704; Conservative 0; Mismatches 463; Indels 156; Gaps 9;

```

```

Qy 60 ATGAAAAAAGCTATTCGCTTATTTTAAACCGCTGATTCATAGGATAGTA 119
Db 1 ATGAAAAAAGCTATTCGCTTATTTTAAACCGCTGATTCATAGGATAGTA 60
Qy 120 TCGAACGCGGGCTGCTACACCTATTTGGATGATTAACCAATATGCTATTT 179
Db 61 TCGAACGCGGGCTGCTACACCTATTTGGATGATTAACCAATATGCTATTT 120
Qy 180 GCCAGATTAAGAGGAAATTCAGAGTGGGCTCAAAATATTAAGTTATTAACAA 239
Db 121 GCCAGATTAAGAGGAAATTCAGAGTGGGCTCAAAATATTAAGTTATTAACAA 180
Qy 240 GGGCAATAGTGGCAATCATGACAAAGCCGATGATGATTTTCTGATGCTCA 299
Db 181 GGGCAATAGTGGCAATCATGACAAAGCCGATGATGATTTTCTGATGCTCA 240
Qy 300 CGTAAAGGCGCTGGCAGCTTGTGAAATCAATATGTTGGAGCGTGCACATACGTA 359
Db 241 AGAAATGAGTGTGCTGCTTATGAGCGATCGTATATGTTGGAGCGTGCACATACGTA 300
Qy 360 GGAATATCAGATGTTGATTTGGTGCAGAGGAAACAACCCGATCAATGCTTTTACT 419
Db 301 GGGTATACCAATGTTGATTTGGTGCAGAGGAAACAATCTGATCAATGCTTTTACT 360
Qy 420 TATAGATTGTAAAGCAATATCTCAAAAGATTAATTTATATCTTTTGAAGAGAT 479
Db 361 TATAGATTGTAAAGCAATATCTCAAAAGATTAATTTATATCTTTTGAAGAGAT 420
Qy 480 TACCATATCAAGATTAATTAATTCGTTACAGAGCGCTCAATGATATGACTTGC 539
Db 421 TACCATATCAAGATTAATTAATTCGTTACAGAGCGCTCAATGATATGACTTGC 480
Qy 540 AATATGATGCGAGTACTTATGATTAAGCAAAATATTCAGAGCGTTCGATGCGC 599
Db 481 CATATGATGCGAGTACTTATGATTAAGCAAAATATTCAGAGCGTTCGATGCGC 540
Qy 600 TCGAAGCGCGCTTTGGCGAATATGATTAAGCAAAAGCGCAAGTTCGGTGCATAT 659
Db 541 TCGAAGCGCGTATTTGGATGATGATTAAGCAAAAGCGCAAGTTCGGTGCATAT 600
Qy 660 CATATGAGAGCTGCAATACACCAATCAGCGTGGAGAGTATGATTCGGTAT 719
Db 600 CATATGAGAGCTGCAATACACCAATCAGCGTGGAGAGTATGATTCGGTAT 719

```

Db 601 AATTATTACAGCTGGGAATACATATATTAAGGCTAGAGGTGATATTCAT 660
 QY 720 TTGGAGGCGAGTGTCTGTAAGCGGAGATATGTCTATACCGATGGAGGCTCAAG 779
 Db 661 GTGAGAGGTGATATGGCAAGGTGGGATTAATGTCTATACCGATGGAGGCTCA 720
 QY 780 GGGGACATGGTCTCTCCGATGTTTATTATGATGCTGAAAAAAGAAATGTTATAT 839
 Db 721 GGGGACATGGATCTCCGATGTTTATTATGATGCTGAAAAAAGAAATGTTATAT 780
 QY 840 GGGATATTACGGGAAGGCAACCTTTGAAAGGCAAAAGAAATGTTGATGTTGCTGC 899
 Db 781 GAGATATGGGAGGAGGCGCAACCTTATACGGGAAATGCGATGATTTCAATGACCG 840
 QY 900 AATCTTTATTTATGATAAATTTTGAAAGAGATTTACATACATCACTTACACCGAG 959
 Db 841 AATCTTTCTCTGATGATAATATACGAAAGATCAACAAATGTTTATACCTTAG 900
 QY 960 GGTATATGAGTATCAATTTATGATGATAATGATGATGATGATGATGATGATGAT 1019
 Db 901 GGGATATGCGTTTATACATTTCTTAAAGAGATGAGGATGAGGATGATGATGATG 960
 QY 1020 TCGAATACCATCAAGAAATTAATTAAGCTGATGATGATGATGATGATGATGAT 1079
 Db 961 ATGGAAAGCC-----TGTGAAATACCTTTAGCGAAACAAATTAAGAGAT 1014
 QY 1080 AAGATTAAGTATATCTATATGATGATGATGATGATGATGATGATGATGATGAT 1139
 Db 1015 AAGATATCTCTAT 1074
 QY 1140 AATGAGAAATGCTAT 1199
 Db 1075 AATGAGAAAT 1134
 QY 1200 ATTAACCAAGGCGGCTGCTTATATTTGAGGATTTTACATGATGATGATGATGAT 1259
 Db 1135 ATGATATGAGTATGAGGCTGCTTATATTTGAGGATTTTACATGATGATGATGAT 1194
 QY 1260 AACCATACTTGGCAAGAGCTGATATATATATATATATATATATATATATATAT 1319
 Db 1195 AATGAAATGAGAAAGGCTGATGATGATGATGATGATGATGATGATGATGATG 1254
 QY 1320 GTAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1379
 Db 1255 GTAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1314
 QY 1380 GCGAAAGGCAAAATTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 1439
 Db 1315 GCGAAAGGCAAAATTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 1374
 QY 1440 CAGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1499
 Db 1375 CAGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1434
 QY 1500 GGGATGCTTCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1559
 Db 1435 GGGATGCTTCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1494
 QY 1560 CGTGTGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1619
 Db 1495 CGTGTGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1554
 QY 1620 GAGGAGGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1679
 Db 1555 GAGGAGGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1614
 QY 1680 GGGAGCAAGAGCAT---TGTCTACCTATGATGATGATGATGATGATGATGATGAT 1735
 Db 1615 GGGAGCAAGAGCAT---TGTCTACCTATGATGATGATGATGATGATGATGATGAT 1674
 QY 1737 AAGAAATGCTTCAACAGGTGCTTGGCAAGATTAATTAATTAATTAATTAATTAAT 1796
 Db 1675 AAGAAATGCTTCAACAGGTGCTTGGCAAGATTAATTAATTAATTAATTAATTAAT 1734

QY 1797 TTAACCTTATTTAATTAACCAACCAAGATGATGATGATGATGATGATGATGATGAT 1856
 Db 1735 TTAACCTTATTTAATTAACCAACCAAGATGATGATGATGATGATGATGATGATGAT 1794
 QY 1857 AATTTAAAGGCAATTAACCAACCAAGATGATGATGATGATGATGATGATGATGAT 1916
 Db 1795 AATTTAAAGGCAATTAACCAACCAAGATGATGATGATGATGATGATGATGATGAT 1854
 QY 1917 CCGACGCTTCAATTAATTAACCAAGATGATGATGATGATGATGATGATGATGATGAT 1976
 Db 1855 CCGACGCTTCAATTAATTAACCAAGATGATGATGATGATGATGATGATGATGATGAT 1914
 QY 1977 GAAATGCTGAGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2036
 Db 1915 GAAATGCTGAGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1974
 QY 2037 AAGGCGAAGTGGGCTGCTTCAATTAACCAAGATGATGATGATGATGATGATGATGAT 2096
 Db 1975 AAGGCGAAGTGGGCTGCTTCAATTAACCAAGATGATGATGATGATGATGATGATGAT 2034
 QY 2097 AGCATATATGCAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2156
 Db 2035 AGCATATATGCAATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2094
 QY 2157 CGTTCAGTGGAGATTAACCAAGATGATGATGATGATGATGATGATGATGATGATGAT 2216
 Db 2095 CGTTCAGTGGAGATTAACCAAGATGATGATGATGATGATGATGATGATGATGATGAT 2154
 QY 2217 ATTAATTCATTAACCAACCAAGATGATGATGATGATGATGATGATGATGATGATGAT 2276
 Db 2155 ATTAATTCATTAACCAACCAAGATGATGATGATGATGATGATGATGATGATGATGAT 2214
 QY 2277 GCGATGTTAAAGTTTAGCAAACTTAATGATGATGATGATGATGATGATGATGATGAT 2336
 Db 2215 GCGATGTTAAAGTTTAGCAAACTTAATGATGATGATGATGATGATGATGATGATGAT 2274
 QY 2337 TTACCTTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2396
 Db 2275 TTACCTTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2334
 QY 2397 -----GCA 2396
 Db 2335 GCAAGGTAATTAATGCACTTATGCGGATGATGATGATGATGATGATGATGATGAT 2394
 QY 2397 -----GCA 2399
 Db 2395 ACATTAAGCAATTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2454
 QY 2400 ACATTAAGCAATTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2459
 Db 2455 ACATTAAGCAATTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2414
 QY 2460 TTAATAAGCAATTTTGGCAAACTTAATGATGATGATGATGATGATGATGATGAT 2519
 Db 2515 TTAATAAGCAATTTTGGCAAACTTAATGATGATGATGATGATGATGATGATGAT 2474
 QY 2520 GAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2579
 Db 2515 GAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2474
 QY 2580 AGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2636
 Db 2635 AGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2594
 QY 2637 CGTTCATTAAGCAATTAACCAAGATGATGATGATGATGATGATGATGATGATGATGAT 2696
 Db 2635 CGTTCATTAAGCAATTAACCAAGATGATGATGATGATGATGATGATGATGATGATGAT 2654
 QY 2697 GTAAATGCTTAAATGAGGCAAGCAATTCATTAATTAATTAATTAATTAATTAAT 2756
 Db 2755 GTAAATGCTTAAATGAGGCAAGCAATTCATTAATTAATTAATTAATTAATTAAT 2714

QY 2757 AAAAGCGATTAATTAATTCATGACGCTGAGGGCGATTATATATCTGTGGC 2816
 DB 2815 GAAAGCGATTAATTAATTCATGACGCTGAGGGCGATTATATATCTGTGGC 2874
 QY 2817 AACAGCGCAAGAACCCGCAACCTTGAAGCAATTAATCTTGTGAAAGCAAGTAAT 2876
 DB 2875 AACAGCGCAAGAACCCGCAACCTTGAAGCAATTAATCTTGTGAAAGCAAGTAAT 2934
 QY 2877 CAACGCTTATCAGATAGCTCAATTTATCTTGAAGTAATGACGCTTATGAGGTCA 2936
 DB 2935 AACCGTTATCAGCAACTCAAAATTTAATTGAAATGACGCTTATGAGGTCA 2994
 QY 2937 TTACGTTAATTAATTAATGAAATGATGCGAATTCGCTTGAATCAATCCATTAAG 2996
 DB 2995 TTACGTTAATTAATTAATGAAATGATGCGAATTCGCTTGAATCAATCCATTAAG 3054
 QY 2997 CAGAAATTCACATGATTTAGTAAGACAGACAGCAAGCAACATTAAGACCA 3056
 DB 3055 CAGAAATTCACATGATTTAGTAAGACAGACAGCAAGCAACATTAAGACCA 3114
 QY 3057 CAAGTTGACCACTGCTTAAACCAACAGGTGACCAAAAGTGGCTCAAGAGACA 3116
 DB 3115 CAAGTTGACCACTGCTTAAACCAACAGGTGACCAAAAGTGGCTCAAGAGACA 3168
 QY 3117 GCGAGACACGCTTCTGATACCTGCTGATCAAGCCTGTTAAAGCATTAAGAC 3176
 DB 3169 ---AGACGCTGTTGCTGATACCTGCTGATCAAGCCTGTTAAAGCATTAAGAC 3225
 QY 3177 AAACAGCTGACCTGCTGATCAACCAAAAGTAAGCAAAACAAAGAGTGGC 3236
 DB 3226 AAACAGCTGACCTGCTGATCAACCAAAAGTAAGCAAAACAAAGAGTGGC 3285
 QY 3237 TCAGAAAGAGC-----AGCTTTCTGATCCCTGCTGATCAAGCCTGTT---C 3284
 DB 3286 TCAGAAAGAGCCTTGAAGAGAGCCTTTCTGATCCCTGCTGATCAAGCCTGTTAAC 3345
 QY 3285 GCATTGAAGCCGCACTTGAAGTATTGATGACCAAGCAATGCGAAAGATGCTA 3344
 DB 3346 GTATTGAAGCCGCACTTGAAGTATTGATGACCAAGCAATGCGAAAGATGCTA 3405
 QY 3345 GCTGAA---GAGAAAGCGCAAAACCAACGCAAAACAAAGCTGATGACCTTATCA 3401
 DB 3406 GCTCAAGAGAGAGAGCGCAAAACCAACGCAAAACAAAGCTGATGACCTTATCA 3465
 QY 3402 AATAGTGGCTATCAGATTAATCTGCAACAGTAATGATGCTTCTGTTCAAGTAA 3461
 DB 3466 AATAGTGGCTATCAGATTAATCTGCAACAGTAATGATGCTTCTGTTCAAGTAA 3525
 QY 3462 TTAGATGCTCTTTTGTAGATCAAGCAATCTGCGTGAACAAATATGCGACAGAT 3521
 DB 3526 TTGATGCTCTTTTGTAGATCAAGCAATCTGCGTGAACAAATATGCGACAGAT 3585
 QY 3522 AAAAGCGCTATGATTTGATGCTGCTGCTTATCAGACAGCAAAACCAATTAAG 3581
 DB 3586 AAAAGCGCTATGATTTGATGCTGCTGCTTAT---CAGACAGAAACCAATTAAG 3642
 QY 3582 CAATTTGGGGTGGCAAAAGCTTATGCTAATGAGAAATTTGGGCGAGTTTCTGCAATAC 3641
 DB 3643 CAATTTGGGGTGGCAAAAGCTTATGCTAATGAGAAATTTGGGCGAGTTTCTGCAATAC 3702
 QY 3642 CGTTCAATTAATCCTTTGATGACAGGTTAAATATCAGCGACATTAAGATGATG 3701
 DB 3703 CGTTCAATTAATCCTTTGATGACAGGTTAAATATCAGCGACATTAAGATGATG 3762
 QY 3702 GGTGTTTCCCAATATCAATGCGGCAATTTCAATTTGGTGAAGTGGGCGGGAAT 3822
 DB 3763 GGTGTTTCCCAATATCAATGCGGCAATTTCAATTTGGTGAAGTGGGCGGGAAT 3882
 QY 3762 AGTGGAGTAAATGCTGAGAAAGCAAGCGCAAAATTCATCGAAAGGATTAAT 3821
 DB 3823 AGTGGAGTAAATGCTGAGAAAGCAAGCGCAAAATTCATCGAAAGGATTAAT 3882
 QY 3822 GAGCGTAATGCAATATCAATTCGTTTAAAGCAATTTGGGCAATTCAGCCTTATTTGGA 3881

DB 3883 GGTGTAAGCAAGTATACGTTCCGTTAGGCAATGGGATTCAGCCTTATTTGGT 3942
 QY 3882 GTTATGCTATTTATTAAGCGTAAATTAATCAATCTGAGAGAGTGAAGTAAAG 3941
 DB 3943 GTTATGCTATTTATTAAGCGTAAATTAATCAATCTGAGAGAGTGAAGTAAAG 4002
 QY 3942 CTAAGCTTCAATTAATGCTAATTAATGCTGAGTATGATTAATTAATTAATCTG 4001
 DB 4003 CTAAGCTTCAATTAATGCTAATTAATGCTGAGTATGATTAATTAATTAATCTG 4062
 QY 4062 ACAGATTAATTAATGCTAATTAATGCTGAGTATGATTAATTAATTAATCTG 4061
 DB 4063 ACAGATTAATTAATGCTAATTAATGCTGAGTATGATTAATTAATTAATCTG 4122
 QY 4062 AACGTAACAGCGGTAATTAATTAATGCTGAGTATGATTAATTAATTAATCTG 4121
 DB 4123 AACGTAACAGCGGTAATTAATTAATGCTGAGTATGATTAATTAATTAATCTG 4182
 QY 4122 AACGTAACAGCGGTAATTAATTAATGCTGAGTATGATTAATTAATTAATCTG 4181
 DB 4183 AACGTAACAGCGGTAATTAATTAATGCTGAGTATGATTAATTAATTAATCTG 4242
 QY 4182 TCTCAAGCTTCACTGCGCAACAGCAATGCGGTAATTAATTAATTAATCTG 4241
 DB 4243 TCTCAAGCTTCACTGCGCAACAGCAATGCGGTAATTAATTAATTAATCTG 4302
 QY 4242 TAA 4244
 DB 4303 TAA 4305

RESULT 7
 US-10-687-046-8
 ; Sequence 8, Application US/10687046
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RT/DCF/DR
 ; CURRENT APPLICATION NUMBER: US/10/687,046
 ; PRIOR FILING DATE: 2003-10-15
 ; PRIOR APPLICATION NUMBER: US/10/080,505
 ; PRIOR FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-23
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 8
 ; LENGTH: 4305
 ; TYPE: DNA
 ; ORGANISM: Haemophilus influenzae
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1702)..(1702)
 ; OTHER INFORMATION: "n" at position 1702 can be any base.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(4305)
 ; OTHER INFORMATION:
 US-10-687-046-8

Query Match 73.6%; Score 3180.8; DB 53; Length 4305;
 Best Local Similarity 85.7%; Pred. No. 0;
 Matches 3704; Conservative 0; Mismatches 463; Indels 156; Gaps 9;

QY 60 ATGAAAAAAGCTATGCTTAAATTTTAAACGCTGCTTATCAATAGGATAGTA 119
 DB 1 ATGAAAAAAGCTATGCTTAAATTTTAAACGCTGCTTATCAATAGGATAGTA 60
 QY 120 TCGAAGCGTGGCTGCTACATTAATTTGGGATTAATTAATTAATTAATTTG 179

Db 51 TCGACCGTGGGCGAGTCACTTATTTGGGATGCACTCAAAATTAATGATGATTT 120

Qy 180 GCGGAGATAAAGGAGAGTTCAGATTGGGAGCTCAAAATATTAAGTTATTAACAACAA 239

Db 121 GCGGAGATAAAGGAGAGTTCAGATTGGGAGCTCAAAATATTAAGTTATTAACAACAA 180

Qy 240 GGGCATTTAGTTGGCAATCAATGACAAAAGCCCCGATGATGATTTTCTGTAGTGTCA 299

Db 181 GGGCATTTAGTTGGCAATCAATGACAAAAGCCCCGATGATGATTTTCTGTAGTGTCC 240

Qy 300 CGTAAAGGCGTGGGAGCGCTGTGAAAATTCATATTTGTGAGCGTGGCAATPACGTA 359

Db 241 AGAAATGAGAGTTGCTGCTTATGTAAGCGATCACTATTTGTGAGTGTGGCAATPAGTA 300

Qy 360 GGAATATACAGATGTGATTTTGTGTGAGAGGAGAAACACCCGATCAATAATGTTTACT 419

Db 301 GGCATATACAGATGTGATTTTGTGTGAGAGGAGAAATTCCTATPACATGTTTACT 360

Qy 420 TATPAGTTGTAAAACGAAATPACTACAAAAGATATTTACATCCTATGAGACGAT 479

Db 361 TATPAAATGTGAAACGAAATPACTATPACATGACATGCAAGACCGCTACTGATGAC 420

Qy 480 TACCATATACAGATTAATTAATTCGTATCAABAACGGCTTCATTTGATGACTTG 539

Db 421 TACCATATACAGATTAATTAATTCGTATCAABAAGATCGGCAACCATGATATGACTTA 480

Qy 540 AATATGATGAGAGTACTTTCAGATGAGACAAATATACGAACCGTGTGATGACG 599

Db 481 CATATGATGAGCAATAGTATGCAATTAAGAAATATTCGAAACGATGAGCGTGGCA 540

Qy 600 TCTGCAAGCGAGTTTGGCGAAATGATCAAGACAAAGCGACCAAGTGGCGTGCATAT 659

Db 541 TCTGCAATCAAGTATTTGGGATGACATCAAAACACAACTTATTCGAGAGATAT 600

Qy 660 CATATCTGCAAGCGTGCCAATPACACAATCAGGTGTGAGACGATATGATGAT 719

Db 601 AATATTTTACAGGTGGGATPACATPACATCAAGGAGTGAAGTGAATTTCAAT 660

Qy 720 TTGGGAGCGATTTGTGTAAAGGGGAGAAATATGTCTCAATTCACATTCGAGGTCGAAG 779

Db 661 GTGAGAGGTGATTTGGCAAAAGTTGGCAATTAATGTCTCATTCGATTTGCAATTTTC 720

Qy 780 GGGGACATGCTTCTCCGATGTTTATATGATGCTGAAAAACAAAAGGTTAATTAAT 839

Db 721 GGGGACATGATCTCCGATGTTTATATGATGCTGAAAAACAAAAGGCTAATTAAT 780

Qy 840 GGGATATTAACGGAGAGCGCAACCTTTTGAAGGCAAGAAAATGGGTTCAATTGGTGC 899

Db 781 GAGATTTTCGGAGAGCGCGCAACCTTTACAGGCGCAATTCGAGGATTTCAATTTGGCCGT 840

Qy 900 AATCTTATTTTGAAGAAATTTGCAAGAGATTTTACATCACTCACTTACACCGAGGT 959

Db 841 AATCTTCTCTGATGAATATTAATACGAAGAAATCAACAAATGGTTTTTTACCCCTAAG 900

Qy 960 GGAATGAGAGTACCAATAGTGAAGATGATATGATGATGATGATGATGATGATGATGAT 1019

Db 901 GGAATGAGAGTATACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960

Qy 1020 TCAAGATACCATCAAGAAATTAATTAAGTGAAGAAATTAAGTGAAGTGAAGTGAAG 1079

Db 961 ATTTGGAAGAAC-----TGTGAAATACCTTTTGGCAACACAAATTAATTAATGAGAT 1020

Qy 1080 AAGATTAAGTTCATATCTATGATATGATATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1139

Db 1015 AAGATTAAGTTCATATCTATGATATGATATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074

Qy 1140 AATGGAAG 1199

Db 1075 AATGGAAG 1134

Qy 1200 AATTAACAG 1259

Db 1135 ATTCGATCAAGTTCAG 1194

QY	1260	AAACCAACTGGCAAGAGCTGGCATACATGTAAAGTAAATAGCACCGCTACTGGAAA	1319
Db	1195	AATGAAGCGTGGAAAGGATGGCGGCAATTCATGTCAGTGAATTAATGACCGTTACTGGAAA	1254
QY	1320	GTAATATGGGTGGAAACATGATGCACTTTCTTAATATGGTAAAGAAACATTGCACTTCAA	1379
Db	1255	GTAAAGGCGGTGGAAATGATGCACTTTCTTAATATGGTAAAGAAACATTGCACTTCAA	1314
QY	1380	GCCAAAGGGGAAATTAAGCTTCGATCAGCGTGAAGGCGATGTAAGTCACTTTGAGCAG	1439
Db	1315	GCCAAAGGGGAAATTAAGCTTCGATCAGCGTGAAGGCGATGTAAGTCACTTTGAGCAG	1374
QY	1440	CAGACAGACGATCAAGGCAACAAACACCTTTATGTAATATGGCTGGTAAAGCGCAGA	1499
Db	1375	CAGACAGACGATCAAGGCAACAAACACCTTTATGTAATATGGCTGGTAAAGCGCAGA	1434
QY	1500	GCGACTGTTCAATTAAACGATGATAACAAATTGATACGCAATAATTATATTCGCGCTT	1559
Db	1435	GCGACTGTTCAATTAAACGATGATAACAAATTGATACGCAATAATTATATTCGCGCTT	1494
QY	1560	CGATGGTGGCTGCTTAATCTTTAAGGCGATTCACTTAACCTTTAAAGTATCCAAATACG	1619
Db	1495	CGATGGTGGCTGCTTAATCTTTAAGGCGATTCACTTAACCTTTAAAGTATCCAAATACG	1554
QY	1620	GAACGAGGGGAGATGATGTGCAACCAATATAGACTCAAGCGCGTAAATGTCATCTTACT	1679
Db	1555	GAACGAGGGGAGATGATGTGCAACCAATATAGACTCAAGCGCGTAAATATCTACT	1614
QY	1680	GGGAACGAAAGCAT--TGTTCTACTTAATGGAATATTAATAAAGCTGATTAACAGA	1736
Db	1615	GGGAACGAAAGTATATGCTGCTCACTTAATAAATAATTAATTAACGTATTAAGC	1674
QY	1737	AAAGAAATGGCTACACAGGTTGGTTGGGAAACAGATAAATTAACACATGGGCGCA	1796
Db	1675	AAAGAAATGGCTACACAGGCTGGTTGGGAAACAGATAAATTAACATTAATGACAGA	1734
QY	1797	TTAAACCTTATTTATTAACCAACACAGAGATCTGACTTGCATCTTCAGTGGTACGA	1856
Db	1735	TTAAACCTTATTTATTAACCAACACAGAGATCTGACTTGCATCTTCAGTGGTACGA	1794
QY	1857	AATTTAAAGGCGATTTATACCAACAAAGTAACATATTTTCAGCGGTAGACGACA	1916
Db	1795	AACCTTAAAGGCGATTTATCTCAACAAAGGTAAACATTTTTCAGCGGTAGACGACA	1854
QY	1917	CCGACGCCCTACATCATTTAAATTAACCTGGTCAAGATGAGAGGTATACCAAGGC	1976
Db	1855	CCGACGCCCTACATCATTTAAACCAACCTGGTCAAGATGAGAGGTATCCCAAGGC	1914
QY	1977	GAATTTGTGTGGATCAAGATTGGATCAACCGTACATTTAAAGCTGAAACCTTCAATT	2036
Db	1915	GAATTTGTGTGGATTAAGATTGGATTAACCGCAATTTAAAGCTGAAACCTTCAATT	1974
QY	2037	AAAGCGGAAGTGGCGGTGTTCTCCGCAATGTTCTCAATGAGGGAATTTGACAGTC	2096
Db	1975	AAAGCGGAAGTGGCGGTGTTCTCCGCAATGTTCTCAATGAGGGAATTTGACAGTC	2034
QY	2097	AGCAATTAAGCAATGCGCATTTGGTGTGTGCCAATCAACAAATACCATTTGGCAG	2156
Db	2035	AGCAATTAAGCAATGCGCATTTGGTGTGTGCCAATCAACAAATACCATTTGGCAG	2094
QY	2157	CGTCCAGATTTGACAGATTAAAGCATTTGTCCAAATAGTGAATTAAACGATATCAAAAGTT	2216
Db	2095	CGTCCAGATTTGACAGATTAAAGCATTTGTAAACAGTTAACTTAACCGATATCAAAAGTT	2154
QY	2217	ATTATATTATACCAAAAACAAATCAATGGCTCTATATCTTAACGATATGCAAGC	2276
Db	2155	ATTATATTATACCAACCAACAAATTAATGGTCTCTATATCTTAACGATATGCAAGC	2214
QY	2277	GCAGATGTTAAAGTATTAGCAAAACCTTAATGGCAATGTGACTTTAAACAAATACAGCGCA	2336
Db	2215	GTGATATTATAGCTTTAGCAAACTTAATGGTAATGTGACTTTAAATTAACATACAGCA	2274

QY 2337 TTACATTAGCAACAATGCGACCCAAATAGCCAAATTCGACTTCCGCAATTCACCT 2396
 DB 2275 TTACATTAGCAACAATGCGACCCAAATAGCCAAATTCGACTTCCGCAATTCACCT 2394
 QY 2397 ----- 2396
 DB 2335 GCAGAGGTAATATATGCGACTTATATGCGAGATGCAATTTAGCGATACCTAGCCGTTT 2394
 QY 2397 -----GCA 2399
 DB 2395 ACATTAAGCAATCAAGCAACACAGATTGGCAATGAGCTTCATCGACAGCTCAAGCA 2454
 QY 2400 AGCGGTATATGCAAACTTGAAGGTATATGCAATTTAGCGATTCAGCTCAATTTCT 2459
 DB 2455 ACAGTATATGCAAACTTGAAGGTATATGCAATTTAGCGATTCAGCTCAATTTCT 2514
 QY 2460 TTAATAAAGCAATTTTCCGCAAAATTCAGGAGACAAAGGCAACAGATGACGTTG 2519
 DB 2515 TTAATAAAGCAATTTTCCGCAAAATTCAGGAGACAAAGGCAACAGATGACGTTG 2574
 QY 2520 GAAATGCGACTTGAACAATGCTTACGATTAATTCAGATTTTACGCTTAAATAC 2579
 DB 2575 GAAATGCGACTTGAACAATGCTTACGATTAATTCAGATTTTACGCTTAAATAC 2634
 QY 2580 AGTACGATCAGCTTAAATTCAGCTTATTCAGCTTACCTAAACAAATAGCCACGT---CGC 2636
 DB 2635 AGTACGATCAGCTTAAATTCAGCTTATTCAGCTTACCTAAACAAATAGCCACGT---CGC 2694
 QY 2637 CGTTCATTAGACGGAACAAAGCCAACTGCGGCAACATGCTTCAACAGCTTCA 2696
 DB 2695 CGTTCATTAGACGGAACAAAGCCAACTGCGGCAACATGCTTCAACAGCTTCA 2754
 QY 2697 GTAAATGTAAATTAAGTGGGCAAGGCAATTCOAATTTCTCATCTTATTTGGCTAT 2756
 DB 2755 GTAAATGTAAATTAAGTGGGCAAGGCAATTCOAATTTCTCATCTTATTTGGCTAT 2814
 QY 2757 AAAAGGATTAATTAATTAATTCAGTATGAGCTGAGGCGCATTAATATCTGTTGCG 2816
 DB 2815 GAAAGGATTAATTAATTAATTCAGTATGAGCTGAGGCGCATTAATATCTGTTGCG 2874
 QY 2817 AACACAGGCAAGAAACCGGAAACCTTGAAGCAATTAATTTGGTGAAGGCAAGATAT 2876
 DB 2875 AACACAGGCAAGAAACCGGAAACCTTGAAGCAATTAATTTGGTGAAGGCAAGATAT 2934
 QY 2877 CAACCTTATCAGATTAAGCTCAATTTACTTTAGAAATATACCAAGCTTATGAGTGA 2936
 DB 2935 AAACCTTATCAGATTAAGCTCAATTTACTTTAGAAATATACCAAGCTTATGAGTGA 2994
 QY 2937 TTAAGTTAATTAAGTGAAGATGATGCGCAATTCGCTTGAATACCAATTAAGAG 2996
 DB 2995 TTAAGTTAATTAAGTGAAGATGATGCGCAATTCGCTTGAATACCAATTAAGAG 3054
 QY 2997 CAGGAATTCAGCAATGATTTAGTATAGACAGAGCAAGCAAGCAATTTGAAGGCCAA 3056
 DB 3055 CAGGAATTCGCTGCTGATTTAGTATAGACAGAGCAAGCAAGCAATTTGAAGGCCAA 3114
 QY 3057 CAAGTTGAACCGACTGCTTAAACAAACAGGTGACCAAAAGTGGCTCAAGAGACA 3116
 DB 3115 CAAGTTGAACCGACTGCTTAAACAAACAGGTGACCAAAAGTGGCTCAAGAGACA 3168
 QY 3117 GCGAGAGCAGGCTTCCGATACCCGCTGATTAAGGCTGTTAAAGCATTAAGAGCC 3176
 DB 3169 ---AGAGCGGTGTTGCTGATACCCGCTGCTCAAGGCTGTTAAAGCATTAAGAGCC 3225
 QY 3177 AAACAAGCTGAATGACTGCTAAACAAACAAAGTATAGGCAAAACAAAGAGTGGCG 3236
 DB 3226 AAACAAGCTGAATGACTGCTAAACAAACAAAGTATAGGCAAAACAAAGAGTGGCG 3285
 QY 3237 TCAAAAAGAC-----AGTGTTCGATCCCTGCTTGAATCAAGGCTGTT---C 3284
 DB 3286 TCAAAAAGACATTAAGAGAGCGTTCGATACCCGCTGATCAAGGCTTAAAGC 3345
 QY 3285 GCATTAAGAGCGCACTTATGATGATGCCCAAGCAATTCGAAAAAGATGCTCTA 3344

DB 3346 GTATTAGAGCGGCACTTAAGCTTATTAATGCGCAACCGCAACAGAAAAAGAGCTCA 3405
 QY 3345 GCTCA---GAAAGAGCGGCAAAACAGCGCAACCAAAAGACTTATGAGCGCTTATTC 3401
 DB 3406 GCTCAAGAGAGAGAGGAAAGCAACAGCAACCAAAAGACTTATGAGCGCTTATTC 3465
 QY 3402 AATATGCGTTATGCAATTTATCTGCAACAGTAATAGTATGCTTCTGTTCAAGATGA 3461
 DB 3466 AATATGCGTTATGCAATTTATCTGCAACAGTAATAGTATGCTTCTGTTCAAGATGA 3525
 QY 3462 TTAGATGCTCTTTTGTATGATCAAGCAATTCGCGTGAAGCAAAATATGCGCAAGAT 3521
 DB 3526 TTGATGCTCTTTTGTATGATCAAGCAATTCGCGTGAAGCAAAATATGCGCAAGAT 3585
 QY 3522 AAAAGCGCTTATGATTTGATGCGGTTCGCTTATGCGAGCAAGCAAACTTACGT 3581
 DB 3586 AAAAGCGCTTATGATTTGATGCGGTTCGCTTAT---CAGAGAAACGAACTTCCGT 3642
 QY 3582 CAATTTGGGCGTCAAAAGCGCTTATGATGAGCAATTTGGGCGAGTTTCTGCGATAC 3641
 DB 3643 CAATTTGGGCGTCAAAAGCGCTTATGATGAGCAATTTGGGCGAGTTTCTGCGATAC 3702
 QY 3642 CGTTCAGATTAATTCCTTGAATGAAACAGTAAATACAGCGCAATTAAGAGATGATG 3701
 DB 3703 CGTTCAGATTAATTCCTTGAATGAAACAGTAAATACAGCGCAATTAAGAGATGATG 3762
 QY 3702 GGTTCGCCCAATATCAATGCGGCGATTTACAAATTTGGTGAAGGCAAGCGCAAT 3761
 DB 3763 GGTTCGCCCAATATCAATGCGGCGATTTACAAATTTGGTGAAGGCAAGCGCAAT 3822
 QY 3762 AGTGGGCAATTAATGCTGCAAGCAAGCGCAAAATTCATGCAAAAGCAATTAAT 3821
 DB 3823 AGTGGGCAATTAATGCTGCAAGCAAGCGCAAAATTCATGCAAAAGCAATTAAT 3882
 QY 3822 GCGGTGAATGCAATTAATGCTGCAAGCAATGCGCAATTTGCGCAATTTGCGCAAT 3881
 DB 3883 GCGGTGAATGCAATTAATGCTGCAAGCAATGCGCAATTTGCGCAATTTGCGCAAT 3942
 QY 3882 GTTAATGCGCTTTTATGCAAGTGAATTAATCAATGCAAGGAGTGAAGTGAAGAG 3941
 DB 3943 GTTAATGCGCTTTTATGCAAGTGAATTAATCAATGCAAGGAGTGAAGTGAAGAG 4002
 QY 3942 CCGAGCTTGAATTAATGCTGCAATTAATGCGCAATTTGCGCAATTTGCGCAAT 4001
 DB 4003 CCGAGCTTGAATTAATGCTGCAATTAATGCGCAATTTGCGCAATTTGCGCAAT 4062
 QY 4002 ACGATTAATTAATGCGCTTAAAGCTTATTTCTGCAATTAATGCTTCAAAAGCT 4061
 DB 4063 ACGATTAATTAATGCGCTTAAAGCTTATTTCTGCAATTAATGCTTCAAAAGCT 4122
 QY 4062 AACGTAAACAGCGTAAATTCAGCGTGTGCAACAGCAATTTGAGCAATTTGAGCA 4121
 DB 4123 AACGTAAACAGCGTAAATTCAGCGTGTGCAACAGCAATTTGAGCAATTTGAGCA 4182
 QY 4122 AAAGAAGTGAATTAAGGCGAAATTTTCAATTCGAAATTTCCGCTTTTATCTCAAA 4181
 DB 4183 AAAGAAGTGAATTAAGGCGAAATTTTCAATTCGAAATTTCCGCTTTTATCTCAAA 4242
 QY 4182 TCTCAAGTTCAAGCTGCGCAACAGCAAAATGTCGCAATTTGAGGCTATGCTGG 4241
 DB 4243 TCTCAAGTTCAAGCTGCGCAACAGCAAAATGTCGCAATTTGAGGCTATGCTGG 4302
 QY 4242 TAA 4244
 DB 4303 TAA 4305

RESULT 8
 US-08-426-787-1/C
 ; Sequence 1, Application US/08426787
 ; GENERAL INFORMATION:
 ; APPLICANT: Robert D. Fleischmann

Query Match	69.9%	Score 3019.8	DB 8	Length 1830121
Best Local Similarity	83.0%	Pred. No. 0		
Matches 3641	Conservative	2	Mismatches 650	Indels 93
				Gaps 14
QY	1	TCATATGCTGTTTAACTA-GTATTTTAAATCGAAAAATTAATTAATTAATTAATTT	59	
Db	278587	TCATATGCTGTTTAAACAGATATTTTAAACGAAAATTAATTAATTAATTAATTAATTT	278522	
QY	60	ATGAAAAAAACGTAATTCGCTTAATTTTAAACCGCTGCAATTCATGAGGATAGTA	119	
Db	278527	ATGAAAAAAACGTAATTCGCTTAATTTTAAACCGCTGCTGCTCAATGAGGATAGTA	278466	
QY	120	TTCGACACGTCGCTGCTGACACTTATTTGGGATATTAACCATATTAACGTAATTT	179	
Db	278467	TTCACACGCTCGGACAGTCATATTTTGGATATCACTACCAATATTAACGTAATTT	278408	
QY	180	GCCGAGATTAAGGAAAGCTTCAACATGTCGGGCTCAAAATTAAGTTTAAACAAACA	239	
Db	278407	GCCGAGATTAAGGAAAGCTTCAACATGTCGGGCTCAAAATTAAGGTTTAAACAAAGA	278348	
QY	240	GGGCAATTGTTGGCCACATCATGAACAAACCCCGATGATGATTTTCTGAAGTCA	299	
Db	278347	GGGCAATTGTTGGGACATCATGAACAAACCCCGATGATGATTTTCCGGGGTGG	278288	
QY	300	CGTAACGGGCTGGGACGCTTGCTGAAAATCATATTTTGGACGGTGGACATTAACGTA	359	
Db	278287	CGTAACGGGCTGGGACGCAATTAAGAACCGCATCATATTTGTAAGGGGACATTAACGGC	278228	
QY	360	GGATATACGATGTGATTTTGGTGCAGAGGAAAAACCCCGATCAACAATCGTTTACT	419	
Db	278227	GGATATACGATGTGATTTTGGTGCAGAAAGGAAACCCCGATCAACAACCGCTTACT	278168	

QY	420	TTAAAGTCTTAAAGAAATPACTAATAA---AAAGTAATTATCACTCTTATAGAGAC	476
Db	278167	TATCAAAATGTGAAAGAAATPATTATCAAGCTTGGGAGAAAGACATCCTTATGATGA	278108
QY	477	GATTACATATATCAAGATATCATTAATTCGTATACAGAGGGGTCCATATGATATGACT	536
Db	278107	GATTATCATATGCTCTGTTTACATTAATTTGTACTGAAGCTGAACCTGTGGGATGACA	278048
QY	537	TGAAATATGATGGCAGTACTTATTCAGATAGAACAAATATCCAGAGGTGTGATC	536
Db	278047	ACAAATATGATGGAAGAAATATATGCTGATATAGAGAACTTCTGAGCGTATGATATA	277988
QY	597	GGCTGTGAGAGCGAGTTTGGCGAATATGATCAAGCAAGGCG-----	639
Db	277987	GGCTGAGAGCGAGTATATGGCGTACAGATTAAGATGAAAGAAAGATGTACATATTTCA	277928
QY	640	-ACCAATTTGCGGTGATATCATATTCAGACGTGGCAATACACATACAGGTGGA	698
Db	277927	TATTAATCTCTAGGTGACATATGTAATCTTACTGAGAAATATCCATATCTACAGTGGGA	277868
QY	699	GCAAGTATATGATATTTGATATTTGGAGCGCATGTTGGTAAACGGGAGAAATATGGTCCA	758
Db	277867	AATGTATATGTAACAGTCATCTTATGTGTAATGATATGCTTATATCATATTAATGTCCA	277808
QY	759	TTACGATATGCAAGGTCAAAAGGGGACAGTGTCTTCGATGTTATTAATATGATGTGAA	818
Db	277807	TTACCAACGGGTGTCTTAAAGGCGCATATAGCGGTTGCGCATATGTTATTAATATGCGAAG	277748
QY	819	AAACAAAATGGTATATTAATGGAATATACGGAAGGCAACCTTTTGAAGCAAGAA	878
Db	277747	AATAAATATGCTTATTAATATGCTATTAACAACTGGGCATCTTTTGGAGAGGT	277688
QY	879	AATGGTTTCATATGTTGCGCAATCTTATTTGAT-----GAAATTTTC	923
Db	277687	AATGGTTTCAGTTAATATGCTATTAACAACTGGGCATCTTTTGGAGAGGT	277628
QY	924	GAAAGGATTAACATATCATCTTACACCGGAGCTGTAATGAGTGTACACAAATTAAT	983
Db	277627	CCTAGGTTTTCACGCTATATTCCTCCCAATTAATATGACATTAATCTCTTGTGATCAAT	277568
QY	984	GGAATATGATATGCTCAGGGGTCTATAATCCAGAAATACGAATATCATATGAAATTA	1043
Db	277567	AATGATGTATCAGGTAAATTAATCTTAATCTGACCTAGTAAGATGTCTTAAAGCAAA	277508
QY	1044	ATTACGTTAGCAATATGATTTACT-----TTGAAGGAAGGATATAGTCT	1094
Db	277507	TGAGATATGAACTGTGAAGTTATTAATCATCTGTTAATCAACCGCTATAGAACT	277448
QY	1095	AATCTTGAATATGAGACCTTAATATTTATCTTCACGTTTAAACATATGAGAAACGCTA	1154
Db	277447	GTTAAAGCAACACAGGCTATTAATTTTACCAACAGATATGAAATATGAAATAATTT	277388
QY	1155	TATTTATATGATCAAAAACAGGATCATTAATCTTGCGATGTGATTAACAGGGGCG	1214
Db	277387	TACCTTGGCACAACAGAAAGAAACCTTAAATATGAAATATTAATATCAAGGTCT	277328
QY	1215	GGTGTCTTATTTTGAAGGTATTTTACATGTCTCGA---ATTCTAACCAATCTGG	1271
Db	277327	GGTGTATTAATCTTGAAGTATTTTGTGTAAGGCAAGCAAAATATTAATTAATCTGG	277268
QY	1272	CAGAGACTGCAATCATATGTAATGTAATATGCAACCGTATCTTGAAGATTAATGGGTG	1331
Db	277267	CAGAGTACAGGGTATCTATATGACAAATGCAATGTGTAATATGAAATATCAATCT	277208
QY	1332	GAAATATGCACTTCTTAATATTTGTAAGGAATTTTGCAAGTCTTAAACCAAGGGGA	1391
Db	277207	GAAATATATGTTATCTTAATATTTGTAATGSCATTTATTAATCAATGTAAGGAAG	277148
QY	1392	ATTAATATGTTCAATGCGTATGAGCATATGTTAAAGTCAATTTTGAGCAGAGGACAGAT	145-
Db	277147	AATTTGGAAGTTTAAGTGGGTATACGCAAAATGCTATCTAATATCAAGACGATATA	277088

1452 CAAGCAACAAACAGCTTTAGTGAATGGCTTGGTTAGCGGACAGGACCTGTTCA 1511
277087 GGGGTCAGAAAACAGCTTTCAAGAGCTTGGCAATGTAAGCGGTGACAGCAAGTTCCA 277028
1512 TTAAACGATGATAACATTTGATACCGATAAATTTATTTGCGCTTGGTGGTGGC 1571
277027 TTAAATAGTACAGATCAGATGATGATCTTAACATATCTATTTGGATTTTGGTGGTGGC 276968
1572 TTAGTCTTAAAGGGGATTCATTAACCTTTAAAGTATCCAAAATACGACGAGGCGCA 1631
276967 TTAGTCTTAAAGGGGATTCATTAACCTTTAAAGTATCCAAAATACGACGAGGCGCG 276908
1632 ATGATGCAACCAATATACACTCAAGCGGTAAATGCTATTAATCTTATGAGCAAGCA 1691
276907 ATGATGCAACCAATATACACTCAAGCGGTAAATGCTATTAATCTTATGAGCAAGCA 1691
1692 AT--TGTCTACTTAATGAAATATATTAATTAATTAATTAATTAATTAATTAATTAAT 276848
276847 ATTAAGTCTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276788
1749 TACAAAGCTTGGTTGGGGAACAGATTAATAAATTAACAAAGGAGATTAACCTTAT 1808
276787 TACAAAGCTTGGTTGGGGAACAGATTAATAAATTAACAAAGGAGATTAACCTTAT 276728
1809 TATTAACCAACACAGAGATCGTACTTGGTCTTCAAGTGTGTAATTAATTAATTAATTAAT 1868
276727 TATTAACCAACACAGAGATCGTACTTGGTCTTCAAGTGTGTAATTAATTAATTAATTAAT 276668
1869 GATTTATCCCAACAAAGGTAACATTTTTCAGCGGTGAGCCGACCGGACGCTTAC 1928
276667 GATTTATCCCAACAAAGGTAACATTTTTCAGCGGTGAGCCGACCGGACGCTTAC 276608
1929 ATCATTAAATTAACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1988
276607 ATCATTAAATTAACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 276548
1989 GATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2048
276547 GATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 276488
2049 GCGGTGCTTTCGCAATGTTTCTCAATGAGGGAATGAGAGTCAAGTCAATTAATGCA 2108
276487 GCGGTGCTTTCGCAATGTTTCTCAATGAGGGAATGAGAGTCAAGTCAATTAATGCA 276428
2109 AATGCCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2168
276427 AATGCCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276368
2169 ACAGATTTACGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2228
276367 ACAGATTTACGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276308
2229 CCAAAAACAAATCAATGCTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2288
276307 CCAAAAACAAATCAATGCTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276248
2289 GGTTAAGCAAACTTAATGCAATGCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2348
276247 GGTTAAGCAAACTTAATGCAATGCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276188
2349 AACATGCAACCCCAATAGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTT 2408
276187 AACATGCAACCCCAATAGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTT 276128
2409 AATGCAAACTTGAACGTTAATGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTT 2468
276127 AATGCAAACTTGAACGTTAATGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTT 276068
2469 AGCAATTTTGCACCAATTTGAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGAC 2528
276067 AGCAATTTTGCACCAATTTGAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGAC 276008
2529 ACTTGACATGCTTGAAGATATCAATGCAATTTAAGCTTAATTAACGTTAATTAACGTTAATTAACGTTAATTAACGTTAAT 2588

276007 ACTTGACATGCTTGAAGATATCAATGCAATTTAAGCTTAATTAACGTTAATTAACGTTAATTAACGTTAATTAACGTTAAT 275948
2589 AGTTAATTAACGTTAATTAACGTTAATTAACGTTAATTAACGTTAATTAACGTTAATTAACGTTAATTAACGTTAATTAACGTTAAT 2645
275947 AGTTAATTAACGTTAATTAACGTTAATTAACGTTAATTAACGTTAATTAACGTTAATTAACGTTAATTAACGTTAATTAACGTTAAT 275888
2646 GAGACGAAACCAAGGCAATGCGCAAGATGCGTTCACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2705
275887 GAGACGAAACCAAGGCAATGCGCAAGATGCGTTCACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275828
2706 AAATGAGTGGGCAAGGACATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2765
275827 AAATGAGTGGGCAAGGACATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275768
2766 AAATTAAT 2825
275767 AAATTAAT 275708
2826 AAAGAACCCGAAACCCCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2885
275707 AAAGAACCCGAAACCCCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275648
2886 TCAGATTAAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2945
275647 TCAGATTAAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275588
2946 AAATTAAGTCAATTAAT 3005
275587 AAATTAAGTCAATTAAT 275528
3006 CACATTAAT 3065
275527 CACATTAAT 275468
3066 CCGATGCTTAAACCAACACAGTGAAGCAAGGAG 3125
275467 CCGATGCTTAAACCAACACAGTGAAGCAAGGAG 275417
3126 GCGTTTCTTAATTCCTGCTGATCAAGGCTGTTAAAGCTTGAAGGCAAGCAAGCT 3185
275416 GCGTTTCTTAATTCCTGCTGATCAAGGCTGTTAAAGCTTGAAGGCAAGCAAGCT 275357
3186 GAATGATGCTGGAAGCAAG 3245
275356 GAATGATGCTGGAAGCAAG 275297
3246 -----GCAATGTTTCTGATCCCTGCTGATCAAGGCTGTTAAAGCTTGAAGGCAAGCAAGCT 3293
275296 GCAATGTTTCTGATCCCTGCTGATCAAGGCTGTTAAAGCTTGAAGGCAAGCAAGCT 275237
3294 GCGGCTTGAAGGTTATGATGCGCAAG 3353
275236 GCGGCTTGAAGGTTATGATGCGCAAG 275177
3354 GAACGGAAGCAAG 3413
275176 G-----AACAGGCAAG 275126
3414 TCAGATTAATTCGCAACAGTAAATGATGCTTCTGTTGAAGCAAGTAAATGATGCTT 3473
275125 TCAGATTAATTCGCAACAGTAAATGATGCTTCTGTTGAAGCAAGTAAATGATGCTT 275066
3474 TTTGATGATCAAGCAATGCTGCGTGAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3533
275065 TTTGATGATCAAGCAATGCTGCGTGAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275006
3534 GATTCGATGCGCTTCTGCTGATCAAG 3593
275005 GATTCGATGCGCTTCTGCTGATCAAG 274949
3594 CAAAAGCTTGAAGTGAAGCAAGTGAAGGCAAGTTCCTGCTGATGAGGCTTCAAGTAAAT 3653


```
Db 274948 CAAAAGCCTTAGATATATGAGCAGATATGAGGAGGCTTTCTGCAATAGCCGTTGAGATAT 274889
QY 3654 ACCCTTGATGAAACAGGTTAAATAATCAGCGAATTAACGATGATGCTGCGTTTGCCAA 3713
Db 274888 ACCCTTGACGAAACAGGTTAAATAATCAGCGAATTAACGATGATGCTGCGTTTGCCAA 274829
QY 3714 TATCATAGGAGGAGATTTACATTTGCTGTAACGTGGAAGGAAATCAGTCCAGATAA 3773
Db 274828 TATCATAGGAGGAGATTTACATTTGCTGTAACGTGGAAGGAAATCAGTCCAGATAA 274769
QY 3774 ATGGCTGAAAGAAACAAAGCCGAAATTCATCGAAACGATTAATTAATGCGTGAATCA 3833
Db 274768 ATGGCTGAAAGAAACAAAGCCGAAATTCATCGAAACGATTAATTAATGCGTGAATCA 274709
QY 3834 AGTTATCAGTCCGTTTAGGGGCAATGGGCAATGAGCTTATTTTGGAGTAAATGCTAT 3893
Db 274708 AGTTATCAGTCCGTTTAGGGGCAATGGGCAATGAGCTTATTTTGGAGTAAATGCTAT 274649
QY 3894 TTTATTTGAAAGTGAATTTATCATCTGAGAGAGTGAAGTGAAGAGCCCTTACCTTGCA 3953
Db 274648 TTTATTTGAAAGTGAATTTATCATCTGAGAGAGTGAAGTGAAGAGCCCTTACCTTGCA 274589
QY 3954 TTTATTCGCTATATCTGCGAATTTGAGTGAATTAATTAATCTCGACAGATTAATATC 4013
Db 274588 TTTATTCGCTATATCTGCGAATTTGAGTGAATTAATTAATCTCGACAGATTAATATC 274529
QY 4014 AGCGTTAGCCTTATTTCTTCTGCGAATTTGAGTGAATTAATTAATCTCGACAGATTAAT 4073
Db 274528 AGCGTTAGCCTTATTTCTTCTGCGAATTTGAGTGAATTAATTAATCTCGACAGATTAAT 274469
QY 4074 AGCGTTAATCTCAGCGTGTGCAACACATTTGAGCGTTATGAGCAAAAGAAAGAGTGGGA 4133
Db 274468 ACTGTAATAGCAGATGTTGCAACATTTGAGCGTTATGAGCAAAAGAAAGAGTGGGA 274409
QY 4134 TTAAGGCAAGAAATTTATGATTTCCAAATTTCCGCTTATTAATCTCAAGGTTTCG 274349
Db 274408 TTAAGGCAAGAAATTTATGATTTCCAAATTTCCGCTTATTAATCTCAAGGTTTCG 274349
QY 4194 CAAGTCGCAAAACAGCAAAATGAGGCGTGAATTTGAGCGTTATGAGCGTTATGAGCGTT 4253
Db 274348 CAAGTCGCAAAACAGCAAAATGAGGCGTGAATTTGAGCGTTATGAGCGTTATGAGCGTT 274289
QY 4254 TAATTTATGCTTATATGATTAACAGAGTGGTCAAGATCCACCTTTTATATTC 4313
Db 274288 TAATTTATGCTTATATGATTAACAGAGTGGG-----GCAATCCACCTTTTATATTC 274234
QY 4314 AATAT 4319
Db 274233 AATAT 274228

RESULT 9
US-10-158-865-1/c
Sequence 1, Application US/10158865
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
FILE REFERENCE: PRI86P2CID1
CURRENT APPLICATION NUMBER: US/10/158,865
PRIOR FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US 05/557,884
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 08/476,102
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: Patenlic version 3.1
SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA
ORGANISM: Haemophilus influenzae
```

```
FEATURE:
NAME/KEY: misc feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (29258)..(29258)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (4416)..(4416)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
```

```

1 LOCATION: (140398)..(140398)
2 OTHER INFORMATION: n equals a,t,c, or g
3 FEATURE:
4 NAME/KEY: misc_feature
5 LOCATION: (.42750)..(.42750)
6 OTHER INFORMATION: n equals a,t,c, or g
7 FEATURE:
8 NAME/KEY: misc_feature
9 LOCATION: (145058)..(145058)
10 OTHER INFORMATION: n equals a,t,c, or g
11 FEATURE:
12 NAME/KEY: misc_feature
13 LOCATION: (145171)..(145171)
14 OTHER INFORMATION: n equals a,t,c, or g
15 FEATURE:
16 NAME/KEY: misc_feature
17 LOCATION: (145942)..(145942)
18 OTHER INFORMATION: n equals a,t,c, or g
19 FEATURE:
20 NAME/KEY: misc_feature
21 LOCATION: (147197)..(147197)
22 OTHER INFORMATION: n equals a,t,c, or g
23 FEATURE:
24 NAME/KEY: misc_feature
25 LOCATION: (150841)..(150841)
26 OTHER INFORMATION: n equals a,t,c, or g
27 FEATURE:
28 NAME/KEY: misc_feature
29 LOCATION: (152500)..(152500)
30 OTHER INFORMATION: n equals a,t,c, or g
31 FEATURE:
32 NAME/KEY: misc_feature
33 LOCATION: (152510)..(152510)
34 OTHER INFORMATION: n equals a,t,c, or g

```

Best Local Similarity 83.0%; Pred. No. 0;
Matches 3641; Conservative 2; Mismatches 650; Gaps 14.

Qy	1	TCGATACGTGCTAACGA - GATATTTTAAATACGAAAAATTACTAATTAATTAACATT	59
Db	278587	TCGATACGTGCTTAACCGATATTTTAAATACGAAAAATTACTAATTAATTAACATT	2785288
Qy	60	ATGAAAAAAACGTATTTTCGTCTTAATTTTAAACCGCTGCAATTCATAGGAGATGA	119
Db	278527	ATGAAAAAAACGTATTTTCGTCTTAATTTTAAACCGCTTGCTGTTCAATAGGAGATGA	2784688
Qy	120	TGCGAACCGTGGCGTGGTCAACCTAATTTTGGAGTTGATTAACCAATATTAACGTATTT	179
Db	278467	TCGACACCGTGGCGGTCAACTTAATTTTGGAGTTGATTAACCAATATTAACGTATTTT	278408
Qy	180	GCCGAGATTAAGGGAGAGTTCAACAGTGGGGCTCAAAATNTAAGGTTATTAACAACAA	239
Db	278407	GCCGAGATTAAGGGAGAGTTCAACAGTGGGGCTCAAAATNTAAGGTTATTAACAACAA	278348
Qy	240	GCGCAATTAAGTGGACATCAATGAACAAAGCCGATGATGATTTTCTGTACTGCA	299
Db	278347	GCGCAATTAAGTGGCAATCAATGAACAAAGCCGATGATGATTTTCTGTACTGCG	278288
Qy	300	CGTAACGGCGTGGCAAGCTTGGTTAAATATATATTGTAGCGTGGCAATTAACGTA	359
Db	278287	CGTAACGGCGTGGCGGATTAATGTAGCGCATATATATTGTAGACCGTGGCAATTAACGCG	278228
Qy	360	GGATATACAGATGTTGATTTTGGTCCAGAGAGGAAACAACCCCGATCAACATGTTTAACT	419
Db	278227	GGATATACAGATGTTGATTTTGGTCCAGAGAGGAAACAACCCCGATCAACATGTTTAACT	278168
Qy	420	TATTAAGATTGTAACGAAATTAACAAA --- AAGATTAATTAATCTCTATAGAGAC	476
Db	278167	TATCAAAATTGTAAAGAAATTAATTAATCAAGCTGGGAGAGAAAGCATCTCTATAGAGGA	278108
Qy	477	GATTACACATATCCAGATTATTAATTTGTTACGAGAGGCTCCAAATTGATTAAGTCT	536

Db 278107 GATTATCATATGCTCGTTTACATTAATTGTAAGTGAAGCTGTAACCTGTTGGTAAATGACA 278048
Qy 537 TCGAATATGATATGAGTACTTCTTACATGTAAGCAAAATATTCAGAACTGTGTCTATC 596
Db 278047 ACAATATGATATGAAAAATGTAATGCTGTAAGAGAACTATCTGAGCTGTAAGTAA 277988
Qy 597 GAGCTGAGAGGAGCTTTTGGGAAATGATCAACAAAGGCG----- 639
Db 277987 GGTTCAGAGCTGCTATTTGGGTAACATTAATGAATGAAGAAACGATGATACATGTTCA 277926
Qy 640 -ACCAAGTCCGGTGCATATCATATCTGACAGCTGGCATACACAATTCAGCTGCA 698
Db 277927 TATTAATGCTGAGGTGCATATGCTTATCTGAGAAATACCATCTCAGAGTGA 277868
Qy 699 GCAAGTAATGATATTCGTAATGAGAGGAGTTCGTAAGCGGAGAAATATGCTCA 758
Db 277867 AATGTAATGATGATGATGATCTGATGTAATGTAAGTGAAGCCATATATGATGCA 277808
Qy 759 TTACGATTTGAGGCTCAAGAGGAGGAGAGCTGCTTCGATGTTTATTTATGATGCTAA 818
Db 277807 TTACCAAGGAGTGTCTTAAAGGAGATGAGGAGTTCGCAATGTTATTTATGATGCAAG 277748
Qy 819 AAACAAAATGTTAAATTAATGAGATAATGAGAGGAGCAACCTTTGAGAGGAGAA 878
Db 277747 AAGAAACAAATGCTTAAATGCTGTAATCAACCTGCAATCTTTTTCGAAAGAGT 277688
Qy 879 AATGGGTTCAATTTGCTGCAATCTTATTTGAT-----GAAATTTTC 923
Db 277687 AATGGGTTCAATTTGCTGCAATCTTATTTGAT-----GAAATTTTC 923
Qy 924 GAAGAGATTTACATCATCATTTACACCCGAGCTGTAATGAGAGTACCAATAGT 983
Db 277627 CCGATGTTTTCAGGCTATATTCCTCCCAATTAATGAGCAATATTCCTTGTATCAAT 277568
Qy 984 GGAATATATATGCTGAGGCTCTAATACTCAAAATCAAGATAACATCAGAAATTA 1043
Db 277567 AATGATGTAAGGTAATTAATCTTACAGCTGTAATGAGAGTCTTAAAGCAAA 277508
Qy 1044 AATGATGTAAGGTAATTAATCTTACAGCTGTAATGAGAGTCTTAAAGCAAA 277508
Db 277507 TCAGAGTGAAGATCTGTAAGTATTAATTCATGCTTAAATCAACCTGTAAGCAAT 277448
Qy 1095 AATCTGATATGAGAGCACTAATAATTTATTCAGCTTAAAGAGTAAAGTCAAT 1094
Db 277447 GTTAAAG 277388
Qy 1155 TATTTTGAATCAAAAACAGAGATCATTTATCTGCACTGACATTAACAGAGGAG 1214
Db 277387 TACCTGGAGCAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 277328
Qy 1215 GGTGCTTTTATTTTGGAGGAGTATTTTACAGTATCTCAA--ATTTCACAAACTTGG 1271
Db 277327 GGTGCTTTTATTTTGGAGGAGTATTTTACAGTATCTCAA--ATTTCACAAACTTGG 277268
Qy 1272 CAGAGAGCTGAGATCATATGTAAGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1331
Db 277267 CAGAGAGCTGAGATCATATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 277208
Qy 1332 GAACATGATGATCTTTTAAATTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1391
Db 277207 GAAATGATGCTTTTAAATTTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 277148
Qy 1392 AATTAAGTTCGATCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1451
Db 277147 AATTTAGAGATTTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 277088
Qy 1452 CAGAGAGCAAAAG 1511
Db 277087 GCGGCTCAAAAG 277028
Qy 1512 TTAAGAGATGATTAAGATTTGATATCCATTAATTTATTTGAGGCTTCTGAGTGTGCTGC 1571
Db 277027 TTAAGATGATGATTAAGATTTGATATCCATTAATTTATTTGAGGCTTCTGAGTGTGCTGC 276968
Qy 1572 TTAAATCTTAAAGGAGATTCATTAACCTTTAAACGATATCAAAATACGAGAGAGAGCA 1631
Db 276967 TTAAATCTTAAAGGAGATTCATTAACCTTTAAACGATATCAAAATACGAGAGAGAGCA 276908
Qy 1632 ATGATGTAAG 1691
Db 276907 ATGATGTAAG 276848
Qy 1692 AT---TCTTACCTTAAG 1748
Db 276847 ATTACTCTCATCTTAATTAAG 276788
Qy 1749 TACAGAGTGTGTTGGGAAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1809
Db 276787 TACAGAGTGTGTTGGGAAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276728
Qy 1809 TATAACCAACAG 1868
Db 276727 TATAACCAACAG 276668
Qy 1869 GATTTTACCAACAG 1928
Db 276667 GATTTTACCAACAG 276608
Qy 1929 AATCATTTAAATTAAGTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1988
Db 276607 AATCATTTAAATTAAGTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 276548
Qy 1989 GATCAGATGATCAACAGTATCAATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2048
Db 276547 GATTAAGATGATCAACAGTATCAATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 276488
Qy 2049 GCGGCTGTTTCTCCCAATGTTTCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2108
Db 276487 GCGGCTGTTTCTCCCAATGTTTCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 276428
Qy 2109 AATGCAATTTGCTGTTGAG 2168
Db 276427 AATGCAATTTGCTGTTGAG 276368
Qy 2169 ACAGATTAAG 2228
Db 276367 ACAGATTAAG 276308
Qy 2229 CCAAAACAG 2288
Db 276307 CCAAAACAG 276248
Qy 2289 GGTGAG 2348
Db 276247 GGTGAG 276188
Qy 2349 AACCAATGAG 2408
Db 276187 AACCAATGAG 276128
Qy 2409 AATGAG 2468
Db 276127 AATGAG 276068
Qy 2469 AGCAATTTTTCAG 2528
Db 276067 AGCAATTTTTCAG 276008
Qy 2529 ACTTGAAG 2588
Db 276007 ACTTGAAG 275948
Qy 2589 AGCTTAATTTGAG 2648
Db 275947 AGCTTAATTTGAG 275888

QY 2646 GAGACGAAACAAGCACAATCGGAGAGACATCGTTTGAACATGTAAGTAATGCT 2705
 DB 275887 GAGACGAAACAAGCACAATCGGAGAGACATCGTTTGAACATGTAAGTAATGCT 275828
 QY 2706 AATTTAGTGGGCAAGGCAATTCATTTACTTCTTATTTGGCTTAAAGCAT 2765
 DB 275827 AATTTAGTGGGCAAGGCAATTCATTTACTTCTTATTTGGCTTAAAGCAT 275768
 QY 2766 AATTTAAATTTATCCATATGAGCTGAGGGGATTAACATTTATCTGTTTGCACACAGGC 2825
 DB 275767 AATTTAAATTTATCCATATGAGCTGAGGGGATTAACATTTATCTGTTTGCACACAGGC 275708
 QY 2826 AAAGACCCGAAACCTTTGAGCATTTAACTTTGTTGAAACCAAGATATGACCGTTA 2885
 DB 275707 AAAGACCTTGACCTTTGAGCATTTAACTTTGTTGAAACCAAGATATGACCGTTA 275648
 QY 2886 TCAATAGCTCAATTTACTTTAGAAATGACACGTTGATGACGCTGATTCATTCGTTAT 2945
 DB 275647 TCAATAGCTCAATTTACTTTAGAAATGACACGTTGATGACGCTGATTCATTCGTTAT 275588
 QY 2946 AATTTAGTGAAGATGATGGCAATTCGCTTGATTAACCCATTTAAAGAGAGAAATG 3005
 DB 275587 AATTTAGTGAAGATGATGGCAATTCGCTTGATTAACCCATTTAAAGAGAGAAATG 275528
 QY 3006 CACAATATTTAGTAAAG 3065
 DB 275527 CACAATATTTAGTAAAG 275468
 QY 3066 CCGACTGCTAAACCAACAGAGTGAAGCCAAAGTGGCTCAAGAGAGAGAGAGAGAG 3125
 DB 275467 CCGACTGCTAAACCAACAGAGTGAAGCCAAAGTGGCTCAAGAGAGAGAGAGAGAGAG 275417
 QY 3126 GCGTTCCGATACCTGCTGCTGATCAAGCTGTTAAAGCATTTGAAAGCCAAACAGCT 3185
 DB 275416 GCGTTCCGATACCTGCTGCTGATCAAGCTGTTAAAGCATTTGAAAGCCAAACAGCT 275357
 QY 3186 GAATGATCTGTAAGACACAAACAAAGTGAAGGCAAAACAAACAAAGTGGCTCAAGAG 3245
 DB 275356 GAATGATCTGTAAGACACAAACAAAGTGAAGGCAAAACAAACAAAGTGGCTCAAGAG 275297
 QY 3246 GAGTGTGTTTCTGATCCCTGCTGATCAAGCTGTT---GCAATTAAGA 3293
 DB 275296 GAGTGTGATGAGGAGGCTTCTGATCCCTGCTGATCAAGCTGTTAAAGTGTTCGA 275237
 QY 3294 GCGGACTTGAAGTATGATGATCCCAAGCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 3353
 DB 275236 GCGGACTTGAAGTATGATGATCCCAAGCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 275177
 QY 3354 GAAGCGAAGAAACAAGCAAAACAAGTGAAGGCAAAACAAAGTGGCTCAAGAG 3413
 DB 275176 G-----AACAAGCGAAGAAACAAGTGAAGGCAAAACAAAGTGGCTCAAGAG 275126
 QY 3414 TCAAGATTTATGCAACAGTAATAGTATCTTCTGTTCAAGATGAATAGATGCTT 3473
 DB 275125 TCGAGTGTGCTGCAACATTAATAGTATGTTTCTGTTCAAGATGAATAGATGCTT 275066
 QY 3474 TTTGATATCAAGCAATTCGCTGCTGCAATATGCAACAGATTAAGAGAGAGAG 3533
 DB 275065 TTTGATATCAAGCAATTCGCTGCTGCAATATGCAACAGATTAAGAGAGAGAGAG 275006
 QY 3534 GATTCGATGCTGCTGCTGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3593
 DB 275005 GATTCGATGCTGCTGCTGCTGAT---CAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 274949
 QY 3594 CAAAAGCCTTGAATATGAGAGATTTGGGGAGTTTCTGCTATGAGAGAGAGAGAGAG 3653
 DB 274948 CAAAAGCCTTGAATATGAGAGATTTGGGGAGTTTCTGCTATGAGAGAGAGAGAGAG 274889
 QY 3654 ACCTTGATGAAGAGTTTAAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3713
 DB 274888 ACCTTGATGAAGAGTTTAAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 274829
 QY 3714 TATCAATGAGGAGATTTCAATTTGATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 3773

DB 274828 TATCAATGAGGAGATTTCAATTTGATGATGAGAGAGAGAGAGAGAGAGAGAGAG 274769
 QY 3774 ATGCTGAAGAAACAAGCCGCAAAATTCATGCAAAAGCATTAATATTAATGAGAGAGAG 3833
 DB 274768 ATGCTGAAGAAACAAGCCGCAAAATTCATGCAAAAGCATTAATATTAATGAGAGAGAG 274709
 QY 3834 AGTTATCAATTCCTTTAGGAG 3893
 DB 274708 AGTTATCAATTCCTTTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 274649
 QY 3894 TTTATGAG 3953
 DB 274648 TTTATGAG 274589
 QY 3954 TTTATGCTGAATATGCTGAGATTCGAGTGAATTAATTAATTAATTAATTAATTAATTA 4013
 DB 274588 TTTATGCTGAATATGCTGAGATTCGAGTGAATTAATTAATTAATTAATTAATTAATTA 274529
 QY 4014 AGGTTAGCCTTATTTCTGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 4073
 DB 274528 AGGTTAGCCTTATTTCTGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 274469
 QY 4074 ACAGTAAATCTCAAGGAGTGGCAACACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 4133
 DB 274468 ACAGTAAATCTCAAGGAGTGGCAACACATTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 274409
 QY 4134 TTTAAAGCAGAAATTTTACATTTCCAAATTTGCTTTATCTGAAATCTCAAGAGTCA 4193
 DB 274408 TTTAAAGCAGAAATTTTACATTTCCAAATTTGCTTTATTTGAAATCTCAAGAGTCA 274349
 QY 4194 CAACCTGGCAAAACAGCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4253
 DB 274348 CAACCTGGCAAAACAGCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 274289
 QY 4254 TAAATTTATGCTTTATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4313
 DB 274288 TAAATTTATGCTTTATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 274234
 QY 4314 TAAATAT 4319
 DB 274233 TAAATAT 274228

RESULT 10
 US-10-329-676-1/c
 : Sequence 1, Application US/10329670
 : GENERAL INFORMATION:
 : APPLICANT: Fleischmann et al.
 : TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fra
 : TITLE OF INVENTION: Titerol, and Uses Thereof
 : FILE REFERENCE: P186P1
 : CURRENT APPLICATION NUMBER: US/10/329,670
 : PRIOR FILING DATE: 2002-12-24
 : PRIOR APPLICATION NUMBER: US 09/643,990
 : PRIOR FILING DATE: 2000-08-23
 : PRIOR APPLICATION NUMBER: US 08/487,429
 : PRIOR FILING DATE: 1995-06-07
 : PRIOR APPLICATION NUMBER: US 08/426,787
 : PRIOR FILING DATE: 1995-04-21
 : NUMBER OF SEQ ID NOS: 1
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 1
 : LENGTH: 1830121
 : TYPE: DNA
 : ORGANISM: Haemophilus influenzae
 : FEATURES:
 : NAME/KEY: misc feature
 : LOCATION: (4747)..(4747)
 : OTHER INFORMATION: n equals a, c, g or c
 : FEATURE:
 : NAME/KEY: misc feature
 : LOCATION: (9921)..(9921)

```
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36511)..(36511)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55309)..(55309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
```



```

NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152530)..(152530)
OTHER INFORMATION: n equals a, t, g or c

```

Query Match 69.9%; Score 3019.8; DB 49; Length 1830121;
 Best Local Similarity 83.0%; Pred. No. 0;

Matches 3641; Conservative 2; Mismatches 650; Indels 93; Gaps 14;

```

QY 1 TCATATGCTTTAACTA-GTATTTTAAACGAAATTAATTAATAAATT 53
DB 278587 TCATATGCTTTAAACGATTTTAAACGAAATTAATTAATAAATT 278588
QY 60 ATGAAAAAACTGATTCGCTTAATTTTAAACCGTGCATTTGATAGGAT 119
DB 278527 ATGAAAAAACTGATTCGCTTAATTTTAAACCGTGCATTTGATAGGAT 278468
QY 120 TCGAAGCGTGGCTGTGCACTATTTTGGAGTGTATACCAATTAATTCGATTT 179
DB 278467 TCGAAGCGTGGCGGATGATTTTGGAGTGTATACCAATTAATTCGATTT 278408
QY 180 GCGGAAATTAAGGAGTGTGCAAGTGGGCTGAAATTTAGGTTTAAACAA 239
DB 278407 GCGGAAATTAAGGAGTGTGCAAGTGGGCTGAAATTTAGGTTTAAACAA 278348
QY 240 GGGCAATTTGGGCACTGATGACAAAGCCCGATGATTTTCTGATGTCA 299
DB 278347 GGGCAATTTGGGCACTGATGACAAAGCCCGATGATTTTCTGATGTCA 278288
QY 300 CGTAACGCGTGGAGCTTGTGTAATTAATTAATTTGAGGCGCAATTAAGTA 359
DB 278287 CGTAACGCGTGGAGCTTGTGTAATTAATTAATTTGAGGCGCAATTAAGTA 278228
QY 360 GATATTAAGTGTGATTTTGGTGAAGGAAACACCCGATCAACCTGTTTACT 419
DB 278227 GATATTAAGTGTGATTTTGGTGAAGGAAACACCCGATCAACCTGTTTACT 278168
QY 420 TATAAGTTGTAACGAAATTAACATCAAA--AAAGATTAATTAATTCCTTAAGAGAC 476
DB 278167 TATAAGTTGTAACGAAATTAATTAATTCCTTAAGAGAGACCTTATGATGA 278108
QY 477 GATTAACATTAATTCAGATTAATTAATTCAGAGGCGCTCAATTTGATAGCT 536
DB 278107 GATTAACATTAATTCAGATTAATTAATTCAGAGGCGCTCAATTTGATAGCT 278048
QY 537 TCGAATTAAGTGTGATTTTCAATTAATTAATTAATTAATTAATTAATTAATTA 596
DB 278047 TCGAATTAAGTGTGATTTTCAATTAATTAATTAATTAATTAATTAATTAATTA 277988
QY 597 GGCTGTGAGCGGATTTTGGGAAATGATCAAGAAAGGCG----- 635

```

```

DB 277987 GGCTGTGAGCGGATTTTGGGAAATGATCAAGAAAGGCG----- 277928
QY 640 -ACCAATGCGCGTGCATATCACTGACGCTGCAATCAACATCAAGGTGA 698
DB 277927 TATTAATGCTGAGGATATTCGTTTCTTCTGAGGAAATCCATCTCAAGGTGA 277868
QY 699 GCAATTAAGTGTGATTTTGGAGGCGGATTTGAAAGCGGAGAAATATGTGCA 758
DB 277867 TATTAATGCTGAGGATATTCGTTTCTTCTGAGGAAATCCATCTCAAGGTGA 277808
QY 759 TTAACGATTCAGGCTGAAAGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 818
DB 277807 TTAACGATTCAGGCTGAAAGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 277748
QY 819 AAAAAAATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 878
DB 277747 AAAAAAATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 277688
QY 879 AATGGCTTCAATTTGCTGCAATTTTGAAT-----GAAATTTTC 923
DB 277687 AATGGCTTCAATTTGCTGCAATTTTGAAT-----GAAATTTTC 277628
QY 924 GAAAGATTTAACAATCACTTAACACCGAGCTGTATAGAGTACACATTAAGT 983
DB 277627 CTAAGTGTTCACGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 277568
QY 984 GGAATGTAATGCTGAGGCTGTAATTAATTAATTAATTAATTAATTAATTAATTA 1043
DB 277567 AATGATGTGATGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 277508
QY 1044 ATAGCTTAAGAAATATGATTAATTT-----TTAAAGAGAGATTAATTAAT 1094
DB 277507 TGAAGATGAGAGATGAGTAAATTAATTAATTAATTAATTAATTAATTAATTA 277448
QY 1095 AATCTGATTAATGAGGAGCTTAATTTATTTTCAACGCTTAAATAGGAAAGCTTA 1154
DB 277447 GTTAAGAGAGAGAGAGCTTAATTTATTTTCAACGAGTAAATAGGAAAGCTTA 277388
QY 1155 TATTTATGATCAAAACAGAGATTAATTTTCAACGATGATTAATTAATTAATTAATTA 1214
DB 277387 TACCTTGGCAGCAGAGAAAGAGATTAATTAATTAATTAATTAATTAATTAATTA 277328
QY 1215 GGTGCTTTATTTTGAAGTAAATTTTACATTTTCA--ATTTAAACAACTTGG 1271
DB 277327 GGTGCTTTATTTTGAAGTAAATTTTACATTTTCA--ATTTAAACAACTTGG 277268
QY 1272 CAAGAGCTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1331
DB 277267 CAAGAGCTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 277208
QY 1332 GAACTGATGACTTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1391
DB 277207 GAACTGATGACTTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 277148
QY 1392 AATAAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1451
DB 277147 AATAAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 277088
QY 1452 CAAGGCAACAAACAGCTTTAGTAAATTTGGCTTGTAGCGGAGAGGCTTTCA 1511
DB 277087 GCGGCTGAAACAGCTTTCAAGAGAGTGTGATTAATTAATTAATTAATTAATTAATTA 277028
QY 1512 TTAACGATTAACAAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1571
DB 277027 TTAACGATTAACAAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 276968
QY 1572 TTAATGTTAAGGAGCTTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1631
DB 276967 TTAATGTTAAGGAGCTTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 276908
QY 1632 ATGATGTGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1691

```



```

OY 3834 AGTATCAATTCCTTGAAGGCAATTTGGCATTGACCTTATTTTGGAGTATGCAAT 3893
DB 274708 AGTATCAATTCCTTGAAGGCAATTTGGCATTGACCTTATTTTGGAGTATGCAAT 274649
OY 3894 TTTATTTAAGGTAATAATATTCATCTGAGAAAGTGAGAGGAAAGCCTAGCTTGCA 3953
DB 274648 TTTATTTAAGGTAATAATATTCATCTGAGAAAGTGAGAGGAAAGCCTAGCTTGCA 274589
OY 3954 TTTAATCGCTTAATATGCTGCAATTTGAGTATTAATATTAATCTCGACAGATTAATTC 4013
DB 274588 TTTAATCGCTTAATATGCTGCAATTTGAGTATTAATATTAATCTCGACAGATTAATTC 274529
OY 4014 AGCGTTAAGCTTATTTCTTCTGCAATTAATATGAGTTTCAAGCGTAACGTAACAC 4073
DB 274528 AGCGTTAAGCTTATTTCTTCTGCAATTAATATGAGTTTCAAGCGTAACGTAACAC 274469
OY 4074 AGCGTAATCTCAAGGTTGCAACCAACATTGAGAGTATTTGGCAAAAAGAGTGGCA 4133
DB 274468 AGCGTAATCTCAAGGTTGCAACCAACATTGAGAGTATTTGGCAAAAAGAGTGGCA 274409
OY 4134 TTAAGGCAAAATTTTACATTTTCCAAATTTCCGTTTATCTCAAAATCTCAAGTTCA 4193
DB 274408 TTAAGGCAAAATTTTACATTTTCCAAATTTCCGTTTATCTCAAAATCTCAAGTTCA 274349
OY 4194 CAATCGCAAAAGCAAAATGAGGAGTATTTGGCTATCTGTTGTAATAATCAACA 4253
DB 274348 CAATCGCAAAAGCAAAATGAGGAGTATTTGGCTATCTGTTGTAATAATCAACA 274289
OY 4254 TAAATTTATGTTATTTGTAATAAGAGTGGCTCAATCAAGTCCCACTTTTATTTCC 4313
DB 274288 TAAATTTATGTTATTTGTAATAAGAGTGGCTCAATCAAGTCCCACTTTTATTTCC 274234
OY 4314 AATAAT 4319
DB 274233 AATAAT 274228

RESULT 11
US-10-329-960-1/c
; Sequence 1, Application US/10329960
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10329,960
; PRIOR FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
```

```

; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40810)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45533)..(45533)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
```

```

1  FEATURE:
2  NAME/KEY: misc_feature
3  LOCATION: (145942)..(145942)
4  OTHER INFORMATION: n equals a, t, g or c
5
6  FEATURE:
7  NAME/KEY: misc_feature
8  LOCATION: (147197)..(147197)
9  OTHER INFORMATION: n equals a, t, g or c
10
11 FEATURE:
12 NAME/KEY: misc_feature
13 LOCATION: (150841)..(150841)
14 OTHER INFORMATION: n equals a, t, g or c
15
16 FEATURE:
17 NAME/KEY: misc_feature
18 LOCATION: (152500)..(152500)
19 OTHER INFORMATION: n equals a, t, g or c
20
21 FEATURE:
22 NAME/KEY: misc_feature
23 LOCATION: (152530)..(152530)
24 OTHER INFORMATION: n equals a, t, g or c
25

```

Query Match	69.98%	Score 3019.08	DB 49	Length 1830121
Best Local Similarity	83.0%	Feat. No. 0		
Matches 3641	Conservative	2	Mismatches 650	Indels 93
			Gaps	14
Qy	1	TCGATAGTGGTTAACTA-GTATTTTAAATACGAAATACTTAATTAATTAACATT	59	
Db	278587	TCATATGTGGTTAACACAGATATTTTAAACGAAATTAATTAATTAATTAACATT	278528	
Qy	60	ATGAAAAAACTGTAATTTGGTCTTAATTTTACCGCTGCAATTCATTAGGATGTA	119	
Db	278527	ATGAAAAAACTGTAATTTGGTCTTAATTTTACCGCTGCTGTTTCATTAAGGATGCA	278468	
Qy	120	TGCGACCGCTGGCGGTGCACTACTCTTTGGATTTGATACCAATATATGTAATTT	179	
Db	278467	TCACAGCGCTGGCGGTGCACTACTATTTGGATTTGATACCAATATATGTAATTT	278408	
Qy	160	GCGGAGATPAAGGGAAGTTCACAGTTGGGGCTCAAAATATTTAGTTTAAACAA	239	
Db	278407	GCGGAGATPAAGGGAAGTTCACAGTTGGGGCTCAAAATATTTAGTTTAAACAA	278348	
Qy	240	GGGCAATAGTTCGACATCAATGA CAAAACCCCGCATGATTAATTTTGTAGTGCA	259	
Db	278347	GGGCAATAGTTCGACATCAATGA CAAAACCCCGCATGATTAATTTTGTAGTGCG	278288	
Qy	330	CGTACGGCGTGGCAGCGCTGGTGA AAAATTAATTAATTTGGACGGGACATGA	359	
Db	278287	CGTACGGCGTGGCAGCGCTGGTGA AAAATTAATTAATTTGGACGGGACATGA	278228	
Qy	360	GGATATACGAATGTGATTTTGGTGCACAGGGAACCAACCCGATCAATCGTTTACT	419	
Db	278227	GGATATACGAATGTGATTTTGGTGCACAGGGAACCAACCCGATCAATCGTTTACT	278168	
Qy	420	TCAAGTCTGTAAACGAATACTACAAA---AAGATATTTACATCCTTAAGAGAC	476	
Db	278167	TTCCAAATTTGTAAAGAAATATTTATGACCTTGGGAAGAAAGCATCCTTAAGTGA	278108	
Qy	477	GATTACCAATTCACGATTCATCAATAATTCCTTAACGAACGGCTCCAAATGATAGCT	536	
Db	278107	GATTACCAATTCACGATTCATCAATAATTCCTTAACGAACGGCTCCAAATGATAGCA	278048	
Qy	537	TCGATATGAAATGGAGTACTTATTCAGATGGAACAAATATCCGAAAGCTGTGGTATC	596	
Db	278047	ACAATATGAAATGGAGTACTTATTCAGATGGAACAAATATCCGAAAGCTGTGGTATTA	277988	
Qy	597	GGCTCTGGAAGGCAATTTTGGCAAAATGATCAACAAAGCG-----	633	
Db	277987	GGCTCTGGAAGGCAATTTTGGCAAAATGATCAACAAAGCG-----	277928	
Qy	640	-ACCAATTCGCGGTGATATCAATTTCTGACAGCTGGCAATACACAAATCGAGCTGCA	698	
Db	277927	TTATATGTCTAGGTGATATTCATTTCTTACGCAAGGAATATCCATATCTAGAGTGGCA	277868	


```

Db 275707 AAGAACCTGTGACCCCTTGAGCAATTAATTGATGTAAGACTTGAATAAACCCTTA 275648
QY 2886 TCAGATAGCTCAAAATTTACTTTAGAAATGACACGCTGATGCGAGGTGATTAGCTAT 2345
Db 275647 TCGATTAAGCTCAAAATTTACTTTAGAAATGACACGCTGATGCGAGGTGATTAGCTAT 275588
QY 2946 AATTATGTAAGATGATGCGAAATTCGGCTTGCAATACCCATTAAGAGCGGAATG 3005
Db 275587 AATTATGTAAGATGATGCGAAATTCGGCTTGCAATACCCATTAAGAGCGGAATG 275528
QY 3006 CACATGATTTAGTAAGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 3065
Db 275527 CACATGATTTAGTAAGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 275468
QY 3066 CCGACTGCTAAACCAACAAAGGTAGCCGCAAAAGGCGGTGCAAGAGACGACGACGAC 3125
Db 275467 CCGACTGCTAAACCAACAAAGGTAGCCGCAAAAGGCGGTGCAAGAGACGACGACGAC 275417
QY 3126 GCGTTTCCTGATACCTGCTGATCAAAAGCTGTTAAAGCATTTAGAGCCAAACAGCT 3185
Db 275416 GCGTTTCCTGATACCTGCTGATCAAAAGCTGTTAAAGCATTTAGAGCCAAACAGCT 275357
QY 3186 GACCTGCTGCTGACAAACAAAGGTAGCCGCAAAAGGCGGTGCAAGAGACGACGACGAC 3245
Db 275356 GACCTGCTGCTGACAAACAAAGGTAGCCGCAAAAGGCGGTGCAAGAGACGACGACGAC 275297
QY 3246 -----GCAGTGTTCCTGATACCTGCTGATCAAAAGCTGTTAAAGCATTTAGAGCC 3293
Db 275296 GCAGTGTTCCTGATACCTGCTGATCAAAAGCTGTTAAAGCATTTAGAGCCGCTT 275237
QY 3294 GCGGACCTGAGGTATGATGACCCCAACAGCAATGCGAAAGATGCTTACCTCAAGAA 3353
Db 275236 GCGGACCTGAGGTATGATGACCCCAACAGCAATGCGAAAGATGCTTACCTCAAGAA 275177
QY 3354 GAAGCGGAAAAACAGGCAACAAAGGATTTGATCAGCCGCTTATCAATGTCCTTA 3413
Db 275176 G-----AACGAGGCAACAAAGGATTTGATCAGCCGCTTATCAATGTCCTTA 275126
QY 3414 TCAGATTTATCTGCAACAGTAAATAGTATGCTTCTGTTCAAGATTAAGATGCTT 3473
Db 275125 TCAGATTTATCTGCAACAGTAAATAGTATGCTTCTGTTCAAGATTAAGATGCTT 275066
QY 3474 TTGTGATCAAGCAATCTGCGGTGCAAAATATGCAAGATTAAGAGCTAT 3533
Db 275065 TTGTGATCAAGCAATCTGCGGTGCAAAATATGCAAGATTAAGAGCTAT 275006
QY 3534 GATTCGATCCGTCGCTGATGACGACGACGACGACGACGACGACGACGACGACGACGAC 3593
Db 275005 GATTCGATCCGTCGCTGATGACGACGACGACGACGACGACGACGACGACGACGACGAC 274949
QY 3594 CAAAAAGCTTATGATGACGATGAGGCGGCTTCTGCGATAGCCGTTCAATAT 3653
Db 274948 CAAAAAGCTTATGATGACGATGAGGCGGCTTCTGCGATAGCCGTTCAATAT 274889
QY 3654 ACCCTTATGACGATGAGGCTTAAATCAGCGACATTAACGATGATGTCGGTTTCCCA 3713
Db 274888 ACCCTTATGACGATGAGGCTTAAATCAGCGACATTAACGATGATGTCGGTTTCCCA 274829
QY 3714 TATCAATGCGGCTTAAATCAGCGACATTAACGATGATGTCGGTTTCCCA 3773
Db 274828 TATCAATGCGGCTTAAATCAGCGACATTAACGATGATGTCGGTTTCCCA 274769
QY 3774 ATGGCTGACGACGATGAGGCTTAAATCAGCGACATTAACGATGATGTCGGTTTCCCA 3833
Db 274768 ATGGCTGACGACGATGAGGCTTAAATCAGCGACATTAACGATGATGTCGGTTTCCCA 274709
QY 3834 AGTATGATGCTGCTTATGAGGCTTATGAGGCTTATGAGGCTTATGAGGCTTATGAGGCT 3893
Db 274708 AGTATGATGCTGCTTATGAGGCTTATGAGGCTTATGAGGCTTATGAGGCTTATGAGGCT 274649
QY 3894 TTTATGAGGCTTAAATCAGCGACATTAACGATGATGTCGGTTTCCCA 3953
Db 274648 TTTATGAGGCTTAAATCAGCGACATTAACGATGATGTCGGTTTCCCA 274589

```

```

QY 3954 TTTATGCTATATGCTGCAATTCGATGATTAACATTTACTGCGACGATTAATTC 4013
Db 274588 TTTATGCTATATGCTGCAATTCGATGATTAACATTTACTGCGACGATTAATTC 274529
QY 4014 AGCGTAAAGCTTATTTCTTCTGCAATTAATGATGATGATGATGATGATGATGATGAT 4073
Db 274528 AGCGTAAAGCTTATTTCTTCTGCAATTAATGATGATGATGATGATGATGATGATGAT 274469
QY 4074 ACGTAAATCTCAGGCTTGAACACCATTTGAGAGCTTATGCGCAAAAGAGGCGCA 4133
Db 274468 ACGTAAATCTCAGGCTTGAACACCATTTGAGAGCTTATGCGCAAAAGAGGCGCA 274409
QY 4134 TTAAGGAGAAATTTATCAATTTCCAAATTCGCTTTATATCTCAAAATCTCAAGTTCA 4193
Db 274408 TTAAGGAGAAATTTATCAATTTCCAAATTCGCTTTATATCTCAAAATCTCAAGTTG 274349
QY 4194 CAACCTGCAACGCAAAATGTCGTAATGTCGTAATGTCGTAATGTCGTAATGTCGTAAT 4253
Db 274348 CAACCTGCAACGCAAAATGTCGTAATGTCGTAATGTCGTAATGTCGTAATGTCGTAAT 274289
QY 4254 TATTTATGCTTATGATTAACAGGCTGCTCAATCCAGCTTTTATTC 4313
Db 274288 TATTTATGCTTATGATTAACAGGCTGCTCAATCCAGCTTTTATTC 274234
QY 4314 AATAT 4319
Db 274233 AATAT 274228

RESULT 12
PCT-US03-05226-12
Sequence 12, Application PC/TUS0305226
GENERAL INFORMATION:
APPLICANT: St. Gene, Joseph W.
FILE OF INVENTION: HEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
FILE REFERENCE: A-59941-1/RET/DCF/DHR
CURRENT APPLICATION NUMBER: PCT/US03/05226
CURRENT FILING DATE: 2003-02-18
PRIORITY APPLICATION NUMBER: US 08/236,791
PRIORITY FILING DATE: 1994-10-25
PRIORITY APPLICATION NUMBER: US 09/839,996
PRIORITY FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 5245
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: CDS
LOCATION: (430)..(4740)
OTHER INFORMATION:
PCT-US03-05226-12

Query Match 65.24; Score 2990.4; DB 1; Length 5245;
Best Local Similarity 82.0%; Pred. No. 0;
Matches 3660; Conservative 0; Mismatches 636; Indels 168; Gaps 12;
QY 1 TCATATGCTTTTACAT-GTATTTTATGCAAAATTCCTAATTAATTAACAT 59
Db 370 TCATATGCTTTTACATGATATTTTATGCAAAATTCCTAATTAATTAACAT 429
QY 60 ATGAAAAAACTGATTTGCTTTATTTTAAACGCTTGAATTCATAGGATAGTA 119
Db 430 ATGAAAAAACTGATTTGCTTTATTTTAAACGCTTGAATTCATAGGATAGTA 489
QY 120 TCGCAAGGCTGCTGCTCACTATTTTGGATTTATTAACATATATGCTATTT 179
Db 490 TCGCAAGGCTGCTGCTCACTATTTTGGATTTATTAACATATATGCTATTT 549
QY 180 GCGGAAATTAAGAGCTCACTGCTGCTCAAAATTAATTAAGCTTATTAACACA 239

```

Db 550 GCTGAGATTAAGGAGATTTCAGTTGGGGCTAAAAATGAGTTTATTAACAAAG 609
 Qy 240 GAGCAATTAGTGGACATCAATGACAAAAGCCCGATGATTTATCTGTAGTCA 299
 Db 610 GGGACTTTAGTGGACATCAATGACAAAAGCCCGATGATTTATCTGTAGTCA 609
 Qy 360 GGTAAAGGCTGGACGCTTGGTGAATCAATATTTGAGCGTGGCAATTAAGTA 359
 Db 670 CGAAATGGGGTGGCGCATAGTAGCGGATGATTTATTTGAGTGGCAATTAAGT 729
 Qy 360 GGTATTCAGATTTGATTTGTGTCAGAGGAAAAGCCCGATGACATGCTTTACT 419
 Db 730 GGAATTAATAGCGTTGATTTGTGTCAGAGGAAAAGCCCGATGACATGCTTTACT 789
 Qy 420 TATAGATTTGAAAAGCAATTAATCAAAAAAGATAATTCATCTTTAGAGAGAT 419
 Db 790 TATCAATTTGAAAAGCAATTAATCAAAAAAGATAATTCATCTTTAGAGAGAT 849
 Qy 480 TACCAATATTCAGATTAATCAATTTGTTAAGAGGCGCTCAATGATTAAGCTTG 539
 Db 850 TATCAATATTCAGATTAATCAATTTGTTAAGAGGCGCTCAATGATTAAGCTTG 909
 Qy 540 AATATGAATGGAGTACTTATTCAGATTAAGAGAAAATTCAGAGATTTGATGGC 599
 Db 910 AATATGAATGGAGTACTTATTCAGATTAAGAGAAAATTCAGAGATTTGATGGC 969
 Qy 600 TCTGAGCGGAGTTTGGCGAATGATCAAGAACAAAGGCGACCA-----644
 Db 970 ACAGTGAACATGTTGGAGGAGCTGATGAAGAACAAAGGAGTAAAGTTCAATGG 1029
 Qy 645 GTTGCAGGTCATTCATTTATCTGACAGCTGGCAATACACATGAGTGGAGAGT 704
 Db 1030 CTTCGATGAGCTTATCTGTGAGAAATGAGCGGTAACACATTCACAAAGTGGAGCGGC 1089
 Qy 705 AATGGAATTTGATTTGGAGCGGATGTTGTAAGGAGGAGAAATGATTCATTCG 764
 Db 1090 AAGGAGCGGTAACCTTAAGTGAAGATACCAAAACCAATTAATGAGATCTCTCT 1149
 Qy 765 AATGAGGCTCAAGAGGAGAGTGGTTCCGAGTGTATTTATGATCTGAAAAACA 824
 Db 1150 ACGGATGTTGCTTGGAGATGATGGTTCCCAATGTTTATTAATGATCAATTAACA 1209
 Qy 825 AATGCTTATTAATGAGATTTACGAGAGGCAACCTTTGAGAGCAAAAGAAATGG 884
 Db 1210 AATGCTTATTAATGAGATTTACGAGAGGCAACCTTTGAGAGCAAAAGAAATGG 1269
 Qy 885 TTGCAATGGTGGCAATCTTAATTTGAGAAA-----TTTCAAGAGATTTACACA 941
 Db 1270 TTCAATTAATTAAGAAAATTTGTTTATGATTAATGTTCTTGTAGAAATTTGCTATA 1329
 Qy 942 TCACCTTAACACCGAGCTGTATGAGATGATCAATTAAGTGAATGATTAAGTCA 1001
 Db 1330 ACATTTTATGAGCAGAGATGAGATGATTAATTTATTTATTTATTTATTTATTTAG 1389
 Qy 1002 GGGCTATTAATC-----AGAAATGAGATTCATCAAAATTAATTAATTTAGTTA 1052
 Db 1390 GGTACGGTATCAACCAATGAAAGAGATTAAGCTCAATTTAAAGTCAAGAGGTT 1449
 Qy 1053 GCAAAATGAGTTTACCTTGAAGAGAGATTAAGTGAATGATTAAGTGAATGAGG 1112
 Db 1450 CAGTATTTATGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1509
 Qy 1113 CTAATATTTATTTCTCAAGTTTAACATGAGAGAGAGAGAGAGAGAGAGAGAG 1172
 Db 1510 GTAAATGCTTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1569
 Qy 1173 CAGAGATCAATTAATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAGAG 1232
 Db 1570 ACAGGAGATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1629
 Qy 1233 GGTATTTAGAGATTTCAAAATTTCAACAAATTTGAGAGAGAGAGAGAGAGAG 1292
 Db 1630 GGTATTTAGAGATTTCAAAATTTCAACAAATTTGAGAGAGAGAGAGAGAGAG 1689

Qy 1293 AGTAAATATAGACCGTATCTGGAAGATTAATGGCGTGAACATGATCGACTTTCAAA 1352
 Db 1690 GGTAAAGAGAGATGATCTGATTTGGAAGATTAAGCGGCTGGAACATGATCGCTTTCAAA 1749
 Qy 1353 ATTGTTAAAGAGATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAGAG 1412
 Db 1750 ATTGTTAAAGAGATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAGAG 1809
 Qy 1413 GGTAGATTAAGATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAGAG 1472
 Db 1910 GGTAGATTAAGATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAGAG 1869
 Qy 1473 AGTAAATATAGACCGTATCTGGAAGATTAATGGCGTGAACATGATCGACTTTCAAA 1532
 Db 1870 GGTAAAGAGAGATGATCTGATTTGGAAGATTAAGCGGCTGGAACATGATCGCTTTCAAA 1929
 Qy 1533 GATACGATTAATTTATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAG 1592
 Db 1930 GATACGATTAATTTATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAG 1989
 Qy 1593 TTACCTTTAAAGGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1652
 Db 1990 TTACCTTTAAAGGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2049
 Qy 1653 ACTCAAGCGGATTAATGATCTATTTAGTGGAGAGAGAGAGAGAGAGAGAGAG 1709
 Db 2050 ACTCAAGCGGATTAATGATCTATTTAGTGGAGAGAGAGAGAGAGAGAGAGAG 2109
 Qy 1710 AATATATTAATTAATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAGAG 1769
 Db 2110 AATATATTAATTAATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAGAG 2169
 Qy 1770 AAGATTAATTAATTAATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAG 1829
 Db 2170 AAGATTAATTAATTAATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAG 2229
 Qy 1830 CGTACTTTGATCTTTCAGTGGTGTCAATTTTAAAGGAGATTTACCAAAAGAGT 1889
 Db 2230 CGTACTTTGATCTTTCAGTGGTGTCAATTTTAAAGGAGATTTACCAAAAGAGT 2289
 Qy 1890 AATGCTTTTACGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1949
 Db 2290 AATGCTTTTACGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2343
 Qy 1950 TCAAGATTAAGATTAATTAATTTGATGATGATTAACAGAGAGAGAGAGAG 2009
 Db 2344 TCAAGATTAAGATTAATTAATTTGATGATGATTAACAGAGAGAGAGAGAGAG 2403
 Qy 2010 ACATTAATGCTGAAGATTTCAATTTAAAGGAGAGAGAGAGAGAGAGAGAGAG 2069
 Db 2404 ACATTAATGCTGAAGATTTCAATTTAAAGGAGAGAGAGAGAGAGAGAGAGAGAG 2463
 Qy 2070 TCTTCAATTTGAGAGATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAG 2129
 Db 2464 TCTTCAATTTGAGAGATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAG 2523
 Qy 2130 CCAATTAATGAGATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAGAGAG 2189
 Db 2524 CCAATTAATGAGATTTGATGATGATTAACAGAGAGAGAGAGAGAGAGAGAGAG 2583
 Qy 2190 AAGTGAATTTAAAGGATTTAAAGGATTTAAAGGATTTAAAGGATTTAAAGGAT 2249
 Db 2584 AAGTGAATTTAAAGGATTTAAAGGATTTAAAGGATTTAAAGGATTTAAAGGAT 2643
 Qy 2250 TCTATTAATTTAAAGGATTTAAAGGATTTAAAGGATTTAAAGGATTTAAAGGAT 2309
 Db 2644 TCTATTAATTTAAAGGATTTAAAGGATTTAAAGGATTTAAAGGATTTAAAGGAT 2703
 Qy 2310 AAGTGAATTTAAAGGATTTAAAGGATTTAAAGGATTTAAAGGATTTAAAGGAT 2369
 Db 2704 AAGTGAATTTAAAGGATTTAAAGGATTTAAAGGATTTAAAGGATTTAAAGGAT 2763

QY 2370 AATATGACCTTCCGACATTCACATGCAACGGTGTATATGCAAACTTGAACGGTAT 2429
 DB 2764 AATATGCACTTTCAATATACGCAATATGCAACGGTGTGAACATGCAATTTGAACGGTAT 2823
 QY 2430 GTGCAATTAAGGATTTACCTCAATTTCTTTAAAAAGACGATTTTTCGCAAAAT 2489
 DB 2824 GTGCAATTTAAGGATTTCTCTCAATTTCTTTAAAAAGACGATTTTTCGCAAAAT 2883
 QY 2490 CAGGAGACAAAGGACCAACAGTGAAGTGGAAAAATGCACTTGGACAAATGCTAGCAAT 2549
 DB 2884 CAGGAGACAAAGGACCAACAGTGAAGTGGAAAAATGCACTTGGACAAATGCTAGCAAT 2943
 QY 2550 ACTACATGCGAATTTAAACGCTAAATTAACATGATGATGATGATGATGATGATGATGAT 2609
 DB 2944 GGCATATGCAATTTAAAGCTAAATTAATGATGATGATGATGATGATGATGATGATGAT 3003
 QY 2610 GGTACCTAAACATATAGCAGCT---CGCGTTCTATTAGAGAGGAAACAGCCCAAC 2666
 DB 3004 GGTACCTAAATATATAGCAGCTGACCGCGTTCTATTAGAGAGGAAACAGCCCAAC 3063
 QY 2667 TCGCAGAACATCGTTTCAACATGTCAGTAAATGTAATTTGAGTGGGCAAGGACA 2726
 DB 3064 TCGCAGAACATCGTTTCAACATGTCAGTAAATGTAATTTGAGTGGGCAAGGACA 3123
 QY 2727 TTCATATTTACTTATCTTTA-TTGGCTAATAAGGATTAATTTAAATTTATCAATGAC 2786
 DB 3124 TTCATATTTACTTATCTTTA-TTGGCTAATAAGGATTAATTTAAATTTATCAATGAC 3183
 QY 2787 GGTAGAGGCGATTTACATATTTATCTGTTGCAACAGGAAAGAACCCGAAACCTTAG 2846
 DB 3184 GGTAGAGGCGATTTACATATTTATCTGTTGCAACAGGAAAGAACCCGAAACCTTAG 3243
 QY 2847 CAATTAATCTTGTGTAAGAACAAAGATATCAACGTTATGATTAAGCTCAATCTTACT 2906
 DB 3244 CAATTAATCTTGTGTAAGAACAAAGATATTAACGTTATGATTAAGCTCAATCTTACT 3303
 QY 2907 TTAGAAATATGACAGCTTATGACAGTGTGATTAAGTAAATTTAGTGAAGATGATGAC 2966
 DB 3304 TTAGAAATATGACAGCTTATGACAGTGTGATTAAGTAAATTTAGTGAAGATGATGAC 3363
 QY 2967 GAATTCGGCTGCAATTAACCAATTAATAAGAGAGATTTGCAATGTTAGTGAAGACA 3026
 DB 3364 GAATTCGGCTGCAATTAACCAATTAATAAGAGAGATTTGCAATGTTAGTGAAGACA 3423
 QY 3027 GAGCAGACAGACCAATTAAGACCAACCAATTTGATGATGATGATGATGATGATGATGAT 3086
 DB 3424 GAGCAGACAGACCAATTAAGACCAACCAATTTGATGATGATGATGATGATGATGATGAT 3483
 QY 3087 GGTGAGCCAAAGTGGGTCAAGAAAGACGCG----- 3119
 DB 3484 AGTAAATGCAAGTGGGTCAAGAAAGACGCGTGTCTGATACCTGCTGATCAAGC 3543
 QY 3120 ----- 3119
 DB 3544 CAGTTAGACGATTTAAGACGCAACAGTGAACGATGCTGTAAGAAACAAAAATAG 3603
 QY 3120 -----AGACGCGTTTCTGATACCTGCTGATCAAGCTG 3158
 DB 3604 GCAAAAAAGTGGGTCAAAAAAGACGCGTGTCTGATACCTGCTGATCAAGCTG 3663
 QY 3159 TTAAGCGATTAAGACCAACAGTGAACGATGATGATGATGATGATGATGATGATGATGAT 3218
 DB 3664 TTAAGCGATTAAGACCAACAGTGAACGATGATGATGATGATGATGATGATGATGATGAT 3723
 QY 3219 AAAACAAAAAGTGGGTCAAAAAAGACGCGTGTCTGATACCTGCTGATCAAGC 3278
 DB 3724 AAAACAAAAAGTGGGTCAAAAAAGACGCGTGTCTGATACCTGCTGATCAAGC 3783
 QY 3279 CTGTTCT---GCATTAGACCGCACTTGAAGTATGATGATGATGATGATGATGATGATGAT 3335
 DB 3784 CGTTAAAGGATTTAGAGGATCAATCTGAGGATTTAATGATGATGATGATGATGATGATGAT 3843
 QY 3336 GATGCTAGCTCAAGAAAGGAGGAAACCAACGCAACAAAAAGCTTGAATGACCGCT 3395

DB 3844 GAACTCAAGATCA-----GAGAAACAGCAAAACAAAGACCTGATGACCGT 3894
 QY 3396 TATTCAATATGAGCTTATATGAAATATGTCGAACAGTAAATGATGCTTTCTGCA 3455
 DB 3895 TATTCAATATGAGCTTATATGAAATATGTCGAACAGTAAATGATGCTTTCTGCA 3954
 QY 3456 GATGAATATGATGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3515
 DB 3955 GATGAATATGATGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4014
 QY 3516 CAGATTAAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3575
 DB 4015 CAGATTAAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4071
 QY 3576 TTAAGTCAATTTGGGCTGCAAAAGCTTATGATGATGATGATGATGATGATGATGATGAT 3635
 DB 4072 TTAAGTCAATTTGGGCTGCAAAAGCTTATGATGATGATGATGATGATGATGATGATGAT 4131
 QY 3636 CATAGCGTTGATTAATACCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3695
 DB 4132 CATAGCGTTGATTAATACCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4191
 QY 3696 ATGTCGGTTTGGCCATATCAATGAGGCGATTTAATGATGATGATGATGATGATGATGAT 3755
 DB 4192 ATGTCGGTTTGGCCATATCAATGAGGCGATTTAATGATGATGATGATGATGATGATGAT 4251
 QY 3756 GGAATCAGTGCAGTAAATATGCTGAAAGAACAGCCGAAATTTATGATGAAAGGATTA 3815
 DB 4252 GGAATCAGTGCAGTAAATATGCTGAAAGAACAGCCGAAATTTATGATGAAAGGATTA 4311
 QY 3816 AATATGAGCTGAATGCAATGATTAATGATGATGATGATGATGATGATGATGATGATGAT 3875
 DB 4312 AATATGAGCTGAATGCAATGATTAATGATGATGATGATGATGATGATGATGATGATGAT 4371
 QY 3876 TTTGAGTAAATGCTATTTTATGAAAGCTGAAATTAATGATGATGATGATGATGATGATGAT 3935
 DB 4372 TTTGAGTAAATGCTATTTTATGAAAGCTGAAATTAATGATGATGATGATGATGATGATGAT 4431
 QY 3936 AAAAGCGTACGCTTGAATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTGAT 3995
 DB 4432 AAAAGCGTACGCTTGAATTAATGCTGATTAATGCTGATTAATGCTGATTAATGCTGAT 4491
 QY 3996 ACTCCGACAGTAAATATGAGCTTAAGCTTATTTCTGCTCAATTAATGATGATGATGAT 4055
 DB 4492 ACTCCGACAGTAAATATGAGCTTAAGCTTATTTCTGCTCAATTAATGATGATGATGATGAT 4511
 QY 4056 AAGCTTAAGTCAAAACAGCTTAATTTCAAGGTTGCAACACATTTGACGTTAT 4115
 DB 4512 AAGCTTAAGTCAAAACAGCTTAATTTCAAGGTTGCAACACATTTGACGTTAT 4611
 QY 4116 TGGCAAAAAGAGTGGGCTTAAGGCAAAATTTAATCAATTTGCAATTTGCTTATG 4175
 DB 4612 TGGCAAAAAGAGTGGGCTTAAGGCAAAATTTAATCAATTTGCAATTTGCTTATG 4671
 QY 4176 TCAAAATCTCAAGTTCAACTCGCAACAGCAAAATGCGGCTGAATTTGGGCTAT 4235
 DB 4672 TCAAAATCTCAAGTTCAACTCGCAACAGCAAAATGCGGCTGAATTTGGGCTAT 4731
 QY 4236 CGTTGTAATAATCAACATATTTTCTGTTATTTGATTAACAGGTTGGTCAATGAC 4295
 DB 4732 CGTTGTAATAATCAACATATTTTCTGTTATTTGATTAACAGGTTGGTCAATGAC 4788
 QY 4296 TCCACCTTTTATTCGCAAT 4319
 DB 4789 --CCACCTTTTATTCGCAAT 4810

RESULT 13

US-10-687-046-12

; Sequence 12, Application US/10687346

; GENERAL INFORMATION:

; APPLICANT: St. Gene, Joseph W.

Query Match	69.2%;	Score 2990.4;	DB 53;	Length 5245;
Best Local Similarity	82.0%;	Pred. No. 0;		
Matches 3660;	Conservative 0;	Mismatches 636;	Indels 168;	Gaps 12

QY	1	TCATAGTGTGTTAACA-GTATTTTAAATACGAAAATTAAGTAATTAATTAACAT	59
Db	370	TCATAGTGTGTTAACAAGCAATTTTAAATACGAAAATTAAGTAATTAATTAACAT	429
QY	60	ATGMAAAAAACGTATTTGCTCTAATTTTAAACCGCTGCATTTCAATTAAGGATGTA	119
Db	430	ATGMAAAAAACGTATTTGCTCTGAATTTTAAACCGCTGCATTTCAATTAAGGATGTA	489
QY	120	TCGCAAGCGTGGGCTGGTCACTAATTTTGGATGTAATTAACAAATTAATTCGGAATTT	179
Db	490	TCGCAAGCGTGGGCGAGTCACTAATTTTGGATGTAATTAACAAATTAATTCGGAATTT	549
QY	180	GCCGAGATAAAGGAAAGTTCAAGATTGAGGCTCAAAATATTAAAGTTTATACAAACA	239
Db	550	GCTGAGATAAAGGAAAGTTTCAGTTGAGGCTCAAAATATTAAAGTTTATACAAAGG	609
QY	240	GGGCATTTGTTGGCAACATCAATGACAAAGCCCGAATGATTTTCTGTAGTGTCA	299
Db	610	GGGCACTTATGTTGGCAATCAATGACAAAGCCCGAATGATTTTCTGTAGTGTG	669
QY	300	CGTACGCGCTGGCAAGCCTTGCTGAAAATCAATATTATGAGCGTGGCATATACGTA	359
Db	670	CGAAATGGGGTGGGGGATTAAGTACGGATCACTAATATTGATGATGTGGCACTACGCT	729
QY	360	GGATATACGATGTGATTTTGTGTGCAAGGGAAACAACCCGATCAACATGTTTAACT	419
Db	730	GGATATATACGCTGTGATTTTGTGACAGAAAGTCCAAATCCCAATCGACATGCTTTAACT	789
QY	420	TATATATATGTAAACGAAATATATACAAAAAAAGAAATTTACATCTTATGAGAGCAT	479
Db	790	TATCAATTTGTAAAGAAATATATATATACGAGGCAAGATACCTTATACATGCTGAC	849
QY	480	TACCATATCAAGATTACATTAATGTGTTACAGAAAGCGCTCCAAATGATATGACTTGG	539
Db	850	TATCAATACCTCTGTTTGCACAAAATTTGTCACTGACGCTGACACACAAAGATACAGAC	909
QY	540	AATATGATATGCAATATTAATTCAGATGAGAACAAATATCCAGACGTTGCTGATTCGGC	599
Db	910	AATATGATATGAGAAATTAATTCAGCTGATTTAAGTAATTAATTCGTTGCTGATTCGGC	969
QY	600	TCGGAACGCGAGTTTGTGGAAATGATCAAGCAAAAGCGACGCA-----	644
Db	970	ACAGGTGAACAAATGTGTGAGGACTGATGTAAGAACAAAGCAAGGAAGTAAAGTTATGG	1022
QY	645	GTTCGCCGTGATATATTTCTGACAGCTGGCAATATACACATCAAGGTGTGAGACAGGT	704
Db	1030	CTTGCTGATGCTTATCTGTGAGAAATGACAGGTATACACACATTTCACAAAGGTGAGCGGCT	1083

QY	705	AATGCAATATGCAATTGCGAAGGAGAGTTCGTAAGAGCGGAGAAATAGATGCTAATACG	764
DB	1090	AAGCGACGGTAACTTAAGTGGAGATATCACAAAACCAATATACATGACCTCTCT	1149
QY	765	ATTGCAAGCTGAAGGGGACAGTGGTCTCGATGTTATTATGATGCTGAAAAACA	824
DB	1150	ACGGGTGTTGGTGGAGATAGTGGTCTCCAAAGTGTATTATGATGCAATTAACAA	1209
QY	825	AAATGGTAAATTATGGGATATACGGGAAGCCATCCCTTTGAGCGAAAGAAATGGG	884
DB	1210	AAATGGCTAATTAATGGCGATTTGCAACCTGTAACCTTTCTGGGAGCTGGAAATGA	1269
QY	885	TTTCAATTGGTGGCAAACTTAATTGATGAAA---TTTCCAAAGAGATTTCATACA	941
DB	1270	TTCCATTAATATGAAAAATATGGTTATATGATATGCTTGTGAAAGATTTCCCTAA	1328
QY	942	TCACTTAACCGGACGCTGTATATGAGATGTCACATTAATGTGAAATGATATGTACG	1001
DB	1330	ACATTTTATGAGCCAGAAAGTAAACGGTCATATTCATTTACTTCAATTAATATGAACT	1389
QY	1002	GGGCTATATACTC-----AGAAATAGGAAATACCAACAGAAATTAATAATAGCTA	1052
DB	1390	GGTACGTTACTCAACGAATGAAAAATGATATGCTCAATTTAAAGTCAGAACGGTT	1449
QY	1053	GCAAAATATGATTTACCTTTGAAAGAGAGATTAAGTCAATCTATCATATGACGA	1112
DB	1450	CAGTTATTTATGAGCATTAAGATTAAGAAAAAGATTAAGAACCTGTTATGCTGCAGGTG	1509
QY	1113	CTTAATATTTTCTCTCAAGTTTAAACAATGAGAAACGCTAATTTTATGATCAAAA	1172
DB	1510	GTAATGCTTTAAACCAAGACTAAATATATGTAATAAATTTACTTTGGAGTACAGA	1569
QY	1173	CAAGATCATTAATCTTCGATCTGACATTTAAACAAAGGCGGGTCTCTTAATTGAG	1232
DB	1570	ACAGGAACCTTAACATTTGAAATATATTAATCAAGGTGTGTGTGTGATTTTGA	1629
QY	1233	GGTAAATTTTACATATCTCAAAATTTCTAACAACTTGGCAGAGACTGGCAATCATGTA	1292
DB	1630	GGTAACTTTAGGTATCTTCGAAAAATATATGCACTTTGGCAAGTGTGGAGTCAATGA	1689
QY	1293	AGTGAATATACACCGTACTCTGGAAAGTAAATGGCGTGAACAATGATGCACTTCTAA	1352
DB	1690	GGTGAAGACGTACTGTATCTTGGAAAGTAAACGGCGTGAACAATGATGCACTTCTAA	1749
QY	1353	ATTGGTAAGGAACATTTGACGTTCAAGCCAAAGGGGAAATTAAGTTGATCAGCGTA	1412
DB	1750	ATTGGTAAGGAACGTTGCAATTCAAGCAAAAGGTGAATAACTAGGCTCATATAGGTA	1809
QY	1413	GGGATGTTAAAGTCTTTTGGACACAGCCAGACGATCAAGGCAACAAACAGGCTTT	1472
DB	1810	GGTGAAGGCAAGTCAATTTATATCAACAGCCGATGGAACACCAAAAACAAGGCTTT	1869
QY	1473	AGTGAATTTGGCTGTGTGACGCGACAGGGGACTGTCAATTAAGATGATTAACATTT	1532
DB	1870	AAAGAGTTGGCATATGATGATGATGAGTCAACGTTCACTAAATATGTCAGATCAAGTT	1929
QY	1533	GATACCGATAATTTATTTGGCTTTGCTGTGTGTGCTTGAATCTTAAACGACATCA	1592
DB	1930	GATCTTAACAATTTATTTATTTGCGATTTGCGTGGTGTGCTTAATCTTAAACGACATCA	1989
QY	1593	TTTAACTTTAATGATCCAAATATGAGGACAGGGGCAATATGTTGAACATATATCA	1652
DB	1990	TTTAACTTTAATGATCCAAATATGAGGACAGGGGCAATATGTTGAACATATATCA	2049
QY	1653	ACTCAAGCGTATATGTCATTAATTAAGTGAACGAAGCAT---TGTTCATCTTAATGA	1709
DB	2050	ACTCAAGTGTCTAATTTATCTTAATTAAGTGAACGAAGCAT---TGTTCATCTTAATGA	2109
QY	1710	AAATATTTAATTAATCTGATTAACGAAAAAGAAATGCTCAACAGTGTGTTGTCGAA	1769
DB	2110	AAATATTTAATTAATCTGATTAACGAAAAAGAAATGCTTCATCAACGTTGTGTTGCGAA	2169

Db	4312	AATATGAGGGAATGCAAGTATCAATGCCGTTTAGGGCAATAGGGCAATGACCTTAT	4317
Qy	3876	TTGGAGTTAATCCGTTATTTTATTTAAGCGAAATATATTCATCTGAGAAATGAGAG	3933
Db	4372	TTGGAGTTAATCCGTTATTTTATTTAAGCGAAATATATTCATCTGAGAAAGTAAAGTG	4431
Qy	3936	AAAACGCTAGCCTGCATTTATGCGATATAGTGCGATTCGAGTTGATATATACATT	3995
Db	4432	AAAACGCTAGCCTGCATTTATGCGATATAGTGCGATTCGAGTTGATATATACATT	4491
Qy	3996	ACTCCACGATATATCAGCGTTAAGCTTATTTCTTCGTAATATGTGATGTTCA	4055
Db	4492	ACTCCACGATATATCAGCGTTAAGCTTATTTCTTCGTAATATGTGATGTTCA	4551
Qy	4056	AACGCTAAGTCAACACGCGAATCTACGCGTTGGACAACCATTTGAGGTAT	4115
Db	4552	AACGCTAAGTCAACACGCGAATAGCACGCGTTGGACAACCATTTGAGGTAT	4611
Qy	4116	TGGCAAAAAGAGTGGATTAAAGCGAAGATTTTACATTTCCAAATTTCCGCTTTATC	4175
Db	4612	TGGCAAAAAGAGTGGATTAAAGCGAAGATTTTACATTTCCAAATTTCCGCTTTAT	4671
Qy	4176	TCAAATCTCAAGGTTGACAACCTGGCGAAACAGCAAAATGTGGCGTGAATGGCGTAT	4235
Db	4672	TCAAATCTCAAGGTTGACAACCTGGCGAAACAGCAAAATGTGGCGTGAATGGCGTAT	4731
Qy	4236	CGTTGTAAAAATCAACATATTTTATCGTTATTTGATTAACAGAGTGCGTCAGATCAGA	4295
Db	4732	CGTTGTAAAAATCAACATATTTGATGCTTTATTTGATTAACAGAGTGCGCGACATC---	4788
Qy	4296	TCCACCTTTTATTCGAATAT	4319
Db	4789	--CGACCTTTTATTCGAATAT	4810

```

RESULT 14
PCT-US03-05226-16
Sequence 15, Application PC/TUS0305226
GENERAL INFORMATION:
APPLICANT: St. Gem, Joseph W.
TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
FILE REFERENCE: A-59941-1/RTT/DCF/DRH
CURRENT APPLICATION NUMBER: PCT/US03/05226
CURRENT FILING DATE: 2003-02-18
PRIOR APPLICATION NUMBER: US 08/296,791
PRIOR FILING DATE: 1994-10-25
PRIOR APPLICATION NUMBER: US 09/839,996
PRIOR FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn version 3.1
SEQ ID NO 16
LENGTH: 4828
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: CDS
LOCATION: (313)..(4548)
OTHER INFORMATION:
PCT-US03-05226-16

```

Query Match	68.8%	Score 2970;	DB 1;	Length 4825;
Best Local Similarity	92.3%	Pred. No. 0;		

<i>Qy</i>	1	TCATAGTGGTTACGA-GATTTTTTAAATACGAAAAATACCTAAATTAATTAACAT	59
<i>Db</i>	253	TCATAGTGGTTAACAGGATTTTTTAAATACGAAAAATACCTAAATTAATTAACAT	312
<i>Qy</i>	6	ATGAAAAAATCGATATTCGTCCTAAATTTTAAACGCTGCATTTTCATTAGAGTAGTA	119
<i>Db</i>	313	ATGAAAAAATCGATATTCGTCCTAAATTTTAAACGCTGCATTTTCATTAGAGTAGTA	372
<i>Qy</i>	120	TCGCAAGCGTGGGCTGGCTACACTAAATTTGGGATTGATTAACCAATTAATTATGGATTTT	179

Db	373	TCCGACAGCGTGGGACAGTCACTACTTATTTGGGATTGATCCAAATATATTCGGATTTT	432
Qy	180	GCCGAGATTAAGGGAAGTTCAAGTTGGGGCTCAAAATATTAGGTTTATTAACAACA	239
Db	433	GCGGAAATTAAGGGAAGTTCAAGTTGGGGCTCAAAATATTATGAGTTTAACTAAAT	492
Qy	240	GGGCAATTAGTTGGCAATCAATGACAAAAGCCCGAGATGATTTTCTGTGTGTCG	299
Db	493	GGAAATTTAGTTGGCAATCAATGACAAAAGCCCGAGATGATTTTCTGTGTGTCG	552
Qy	300	CGTACGCGGTGACAGCTTGGTTGAAATCAATATATTGTGAGCGTGGCACTAACGTA	359
Db	553	CGAAATGGGGTGTGGGATTTGGTGGGGGAGTACAGTATATTGTGAGTGTGCACTAATGTA	612
Qy	360	GGATATACAGATGTTGATTTTGGTGGCAGAGGGAAAACAACCCCATCAACATCGTTTACT	419
Db	613	GGCTATACCAATGGGATTTTGGTGGTGGAAAGACAAATTCCTATCAACATCGTTTACT	672
Qy	420	TATAGATGTATTAACGAAATATCTACAAAAGATTAATTACATCCTTATAGGAGAGAT	479
Db	673	TATTAATTTTGGAAACGGAAATATTTATTAAGAGATCAAAAGCATCCTTATAGGAAAGC	732
Qy	480	TACCAATATCCGATTAATGATTAATTCGTTACGAAGGCGTCCAAATGATATGACATTGG	539
Db	733	TACCAACACCCACCTTACATTAATTTGTTACGAAAGCACCCCAATCATATACCTTCT	792
Qy	540	AATATCAATGCGAGTACTTATTCAGATAGAACTCAAGATCTGAGATCGTTCGATCGGC	599
Db	793	GATTTAAACGCAACCAATATACAGATAGAACGAAATATCCCAAGCGGTGCGATATGGC	852
Qy	600	TTGAGACGGCGATTTTGGGGAATGATCAAGCAAAAGCGACAGATTTGGCGGTGCAT	659
Db	853	TCCGGGTGGCGATTTTGGGGAACGATCAAAACCAAGCGCACAAATGTTGGCGGCGATAT	912
Qy	660	CATTATCTGAAAGGTGGCAATACACACAAACGCTGAGAGAGATAGATATTCGAT	719
Db	913	CATTACTCGAAGAGGCAATACACACAAAGCGAGAGAGGAGGCGCGTGGTCAAGT	972
Qy	720	TTGGAGAGCGAATGTTGTGTAACCGGAGATATGCTCCATTAACGATTTGACAGCTCAAG	779
Db	973	CTAGAGGGGATGTGGCCAAACGGGCAATTAAGGCCCATTCCTATATGACGGCTCAAG	1032
Qy	780	GGGGAACGTGTTCTCCATGTTTATTTATGATGCTGAAAAAQAATAATGTTAATTAT	839
Db	1033	GGGGAAGGGTGGCTCTGTTTATTTAGATGGGAAAAAQAATAATGTTAATTATAC	1092
Qy	840	GGGATTTAACGGGAGGCAACCTTTGAAAGCAAAAGATGGGTTTCATTTGTTGCTTGC	899
Db	1093	GGGATTTTGAAGACCGGCAACCTTTGGGGGGGAGAGGAATACATTCCTCAACTGTATGC	1152
Qy	900	AAATCTTATTTTGTAGAAATTTTGAAAGATTTACATCATCATCATTACACCGAGCT	959
Db	1153	AACTCTTTTGTAGAAATTCCTGAAAAAGATTTGGCTCATCGTTTATTAAGCCATCG	1212
Qy	960	GGTAATGAGGTATACAAATATGATGGAATATGATATGCTGAGGGGTATATACTCGAAA	1019
Db	1213	GGGATGAGGTATACCACTTACAGCAAAAGCGAGCGACAGGCAATGTCAAAACAACA	1272
Qy	1020	TCAGAAATPACATCAAAATTAATATAGCTTACGCAATATGAGTTTACCTTTGAAAGAG	1079
Db	1273	ACAGGAAGCATGTAGACTCGGATCGGTTTAAAGACAGCAAAATTAACCGCGCAAGGT	1332
Qy	1080	AAAGATTAAGTTCAATCTTATGATATGACGACCTAATATTATTTCTCACGTTTAAC	1139
Db	1333	AAAGAGATGTTTACCA-----ATACCAAGGTTCCAAAATATATACCTGCTCGTTGAAT	1386
Qy	1140	AATGAGAAAGCTATATTTATGATCAAAAACAAGGATCATTAATCTTGCACTGCAC	1189
Db	1387	AAAGGAGAACTGTTTATGAGATCAAAAAACAAGGATCATTAATCTTATGACAAAC	1446
Qy	1200	ATTTAACCAAGGGGCGGTGCTTTATTTTGAAGGTAATTTACAGTATCTCCAAATCT	1258

Db 1447 ATCAACGAGGTCGGGCGCTTTGTTATTTTGAAGGTAACCTTACGTAATCTTCAGAAAT 1506
 Qy 1260 AACCAAACTTGGCAAGACCTGGATCATAGTAAGTAAATAGCACCGTTACTTGAAA 1319
 Db 1507 AATGCAACTTGGCAAGGTCCTGGATCATAGTAAAGTAAAGACAGTACTTACTTGAAA 1566
 Qy 1320 GTAAATGGCTGGCAAGTATGATGACTTTTAAATTTGTAAGGACATTTGACGTTCA 1379
 Db 1567 GTAAATGGCTTTGAAATGATGCTTTTAAATTTGCAAAAGGCAATTTGACGTTAA 1626
 Qy 1380 GCCAAAGGGGAAAATTAAGTTGATGATGAGGTGAGGTGTAAGTCAATTTTGAAGAG 1439
 Db 1627 GCCAAAGGGGAAAATTAAGTTCATGAGGTGAGGTGTAAGTCAATTTTGAAGAG 1486
 Qy 1440 CAGCAGACATGACAGCAACCAAGCCTTAGTAAATTTGCTTGGTTAGCGGAGA 1499
 Db 1687 CAGCAGACATGACAGCAACCAAGCCTTTAGTAAATTTGCTTGGTTAGCGGAGA 1746
 Qy 1500 GGGACTGTTCAATTAAGCATGATTAACAATTTGATACCAATTAATTTTATTCGGCTTT 1559
 Db 1747 GGTACGGTTCAATTAAGCATGATTAACAATTTGATACCAATTAATTTTATTCGGCTTT 1806
 Qy 1560 CGTGGTGGTGGTAAATCTTAAGGGGATTTCACTTAACTTAAAGTAAAGTAAAGTAAAG 1619
 Db 1807 CGTGGTGGTGGTAAATCTTAAGGGGATTTCACTTAACTTAAAGTAAAGTAAAGTAAAG 1866
 Qy 1620 GACGAGGAGGCAATGATTTGTAACCATATACACTCAAGCGGCTATGCTATTTACT 1679
 Db 1867 GATGAGGAGGCAATGATTTGTAACCATATACACTCAAGCGGCTATGCTATTTACT 1926
 Qy 1680 GGGAGCAAAAGCTTTGTTCTACCTAATGAAATATATTAATTAACCTTGTATTCAGAAA 1739
 Db 1927 GCGAGCAATACCTTATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1977
 Qy 1740 GAAATGGCTACAGCTTGGTTGGGAAAAGTAAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1799
 Db 1978 ---ATTCCTTTATATGTTGGTTGGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2034
 Qy 1809 AACCTTATTTATTAACCAACAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1859
 Db 2035 AATGTTACTTACCAATCCGTTTAAAGAAATTAACCTTCTTATGATAGGTGAAACAAAT 2094
 Qy 1869 TTAAGAGGCAATTTACCAACCAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1919
 Db 2095 TTAAGAGGCAATTTACCAACCAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2154
 Qy 1920 CAGGCTACATCAATTTAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1979
 Db 2155 CAGGCTACATCAATTTAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2208
 Qy 1980 ATTTGTTGGGATCAAGTATGATCAACCGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2039
 Db 2209 GTGGTTATGATCAAGTATGATCAACCGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2268
 Qy 2040 GGGGAGTGGCGGTTTCTCGAATGTTCTCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2099
 Db 2269 GGGGAGTGGCGGTTTCTCGAATGTTCTCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2328
 Qy 2100 AATATGCAATGCCAATTTGTTGTTGTCGAATCAACCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2159
 Db 2329 AATATGCAATGCCAATTTGTTGTTGTCGAATCAACCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2388
 Qy 2160 TCAATTTGCAAGATTAAGCATTTGCAAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2219
 Db 2389 TCAATTTGCAAGATTAAGCATTTGCAAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2448
 Qy 2220 AATTTATACAAAACCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2279
 Db 2449 AATTTATACAAAACCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2508
 Qy 2280 AATTTATACAAAACCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2339
 Db 2509 AATTTATACAAAACCAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2568

Qy 2340 AATTAAGCAACATGCGACCCAAATAGCAATATGCACTTTCCGACATTTCACTGCA 2399
 Db 2569 AATTAAGCAACATGCGACCCAAATAGCAATATGCACTTTCCGACATTTCACTGCA 2628
 Qy 2400 ACGGTGATTAATGCAATCTTAAAGGTAATGCAATTTAAGGATTTGCAATTTTCT 2459
 Db 2629 ACGGTGATTAATGCAATCTTAAAGGTAATGCAATTTAAGGATTTGCAATTTTCT 2688
 Qy 2460 TTAAGAAACAGCAATTTTGGCACCAATTTGAGGAGACAAAGGCAACATGTAAGCTTTG 2519
 Db 2689 TTAAGAAACAGCAATTTTGGCACCAATTTGAGGAGACAAAGGCAACATGTAAGCTTTG 2748
 Qy 2520 GAAATAGCACTTGGACACATGCGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2579
 Db 2749 GAAATAGCACTTGGACACATGCGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2808
 Qy 2580 AGTACATCACTTAAATTTGAGCTTATTCAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2636
 Db 2809 AGTACATCACTTAAATTTGAGCTTATTCAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2868
 Qy 2637 CGTTCTTAAAGAGGCAACCAAGGCAATGCGGAGACATGCGGAGACATGCGGAGACATGCGGAGACATGCGGAG 2696
 Db 2869 CGTTCTTAAAGAGGCAACCAAGGCAATGCGGAGACATGCGGAGACATGCGGAGACATGCGGAGACATGCGGAG 2928
 Qy 2697 GTAAATGTAATTTGAGTGGGCAAGGCAATTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 2756
 Db 2929 GTAAATGTAATTTGAGTGGGCAAGGCAATTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 2988
 Qy 2757 AAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2816
 Db 2989 AAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3048
 Qy 2817 AACAGGCAAGAAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 2876
 Db 3049 AACAGGCAAGAAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 3108
 Qy 2877 CAACCTTATGAT 2936
 Db 3109 CAACCTTATGAT 3168
 Qy 2937 TTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2996
 Db 3169 TTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3228
 Qy 2997 CAGGATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3056
 Db 3229 CAGGATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3288
 Qy 3057 CAAGTTGACGACCTGCTAAACCAACCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3116
 Db 3289 CAAGTTGACGACCTGCTAAACCAACCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3342
 Qy 3117 GCGAGAGCAGGTTTCTGATATCTGCT 3176
 Db 3343 ---AGAGGCGGTTTCTGATATCTGCT 3399
 Qy 3177 AACCAAGCTGAATGATGCTGCTGCAACCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3231
 Db 3400 AACCAAGCTGAATGATGCTGCTGCAACCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3459
 Qy 3460 ACTGACCTTAACCAAGAAAGAGCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3519
 Db 3520 ACTGACCTTAACCAAGAAAGAGCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3579
 Qy 3288 TTAAGAGGCGGCTGATGATTTGATGCT 3321
 Db 3322 TTAAGAGGCGGCTGATGATTTGATGCT 3379
 Qy 3322 ---AGCATGCGAA 3332
 Db 3580 GCTCAATGATTTAACAACAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3635

```

1 CURRENT FILING DATE: 2003-10-15
2 PRIOR APPLICATION NUMBER: US/10/080,505
3 PRIOR FILING DATE: 2002-02-22
4 PRIOR APPLICATION NUMBER: US 08/296,791
5 PRIOR FILING DATE: 1994-10-25
6 PRIOR APPLICATION NUMBER: US 09/839,996
7 PRIOR FILING DATE: 2001-04-20
8 NUMBER OF SEQ ID NOS: 58
9 SOFTWARE: PatentIn version 3.1
10 SEQ ID NO: 16
11 LENGTH: 4828
12 TYPE: DNA
13 ORGANISM: Haemophilus influenzae
14 FEATURE:
15 NAME/KEY: CDS
16 LOCATION: (313)..(4546)
17 OTHER INFORMATION:
18 US-10-687-046-16

```

Query Match	68.8%	Score 2970;	DB 53;	Length 4828;
Best Local Similarity	82.3%;	Pred. No. 0;		
Matches 3581; Conservative	0;	Mismatches 645;	Indels 124;	Gaps 10;

Qy	1	TCATAGTGGTTACGA-GATATTTTAAACGAAATTTCTATTTAAATTAACCT	59
Dp	253	TCATAGTGGTTACGACGATATTTTAAACGAAATTTCTATTTAAATTAACCT	312
Qy	60	ATGAAAAAACTGATTTTCCTTAATTTTAAACGCTTGATTCATTAAGGATAGTA	115
Dp	313	ATGAAAAAACTGATTTTCCTGATATTTTAAACGCTTGATTCATTAAGGATAGTA	372
Qy	120	TCCAAAGGTGGGTGCTCAACTATTTGGGATTGATTCGAAATTAATGCTGATTT	179
Dp	373	TCCAAAGGTGGGTCAAGCTCACTATTTTGGGATTGATTCGAAATTAATGCTGATTT	432
Qy	180	GCCGAGATTTAAAGGGAAGTTCACTAGTTTGGGCTCAAAATATTTAAGGTTTAAACAAACA	239
Dp	433	GCCGAGATTTAAAGGGAAGTTCACTAGTTTGGGCTTAAAAATTTGAGGTTTACGATTAAT	492
Qy	240	GGGCAATTAGTTGGCACATCATATGACAAAGCCCGACGATTAATTTTCTGTAGTCA	299
Dp	493	GGAATTTAGTTGGCACATCATATGACAAAGCCCGACGATTAATTTTCTGTAGTCA	552
Qy	300	CGTAAACGGCGTGGAGCGTTGGTGAATAATCATATTTGGAGCGTGGACATTAAGCTA	359
Dp	553	CGAATGGGGTGGCGGCAATTTGGTGGCGCATCATATTTGGAGTGGGACATTAAGCTA	612
Qy	360	GGATATACAGATGTTATTTTGGTGCAGAGGAGAAACAACCCGATTCACATCGTTTTACT	419
Dp	613	GGCTATACCATATGGATTTTGGTCTGAAAGGCAAAAATCCGATTCACATCGTTTTACT	672
Qy	420	TATAGATTTGTAATCGAATATACATCAAAAAGATTAATTACATCTTTATGAGCAT	479
Dp	673	TATTAATTTGTAATCGGAATATATTAATAAGATCAACGATCTTTATGAGGAATC	732
Qy	480	TACCAATATCACGATTAACATTAATTTGGTTACAGAGGGCTCCATTTGATGACTCG	539
Dp	733	TACCAACATCCAGCTTACATTAATTTGTTACGGAAGCCACCGCATCATATGACTTCT	792
Qy	540	AATATGAATGGCAGTACTTATTCAGATAGAACAAAATATCCAAACGTTGTTGATGGGC	599
Dp	793	GATATGAACGGCAACAAATATACAGATAGAACAAATATCCCAACCGTGGTATCGGC	852
Qy	600	TCGGAACGCGAGTTTGGGGAATGATCAAGACAAAGCGAACAGTGGCGGTGCATAT	659
Dp	853	TCGGGTGGCGTTTGGGGAACGATCAAAACAAAGCGGAACAGTTGTCGGCGCATAT	912
Qy	660	CATTATCTGACAGCTGGCAATACACACATCAGCTGAGAGCGATATGATATATTCGAT	719
Dp	913	CATTACTGACAGCGGCAATACACACACAGCGGAGCAGGAGGCGCTGTGCAAGT	972
Qy	720	TTGGAGGCGCATTTGTAAACGGGAGAAATATGATTCATTCGATTTGACGCTCAAG	779

Db 973 CTGAGGGCGATTGCGCCCAAGGCGCAATTAACGCCCATTCATTGCGAGCTCAAGC 1032
 Qy 780 GGGGACAGTGGTCTCCGATGTTATTTATGATGCTGAAAAAACAATAATGTTAATAT 839
 Db 1033 GGGGACAGGCGTTCCTATGTTATTTATATGCGGAAAAACAATAATGTTGATTAC 1032
 Qy 840 GGGGATTTACGGGAAGCGCAACCTTTTGAAGGCAAAATAGGTTTCAATGGTTGCG 899
 Db 1093 GGGGATTTAGGAGACCGGCAACCTTTGGCGGGGCAAGAGACACATTCCTCACTGGTACGC 1152
 Qy 900 AATCTTATTTGATGAAATTTTCAGAAAGATTACATACATCACTTACACCGAGCT 959
 Db 1153 AAGTCTTTTGTGATGAAATTCCTGAAATAAGATTTGCTGACATGTTTATGACCATG 1212
 Qy 960 GGTAAATGAGTGCACACATTAATGTAATGTAATGTCAGGCTCTAATACCTCAAA 1019
 Db 1213 GGGATGAGTGCATCACATTAACAGAAAGAGGAGCGGAGGATGTCAGAACAGCA 1272
 Qy 1020 TCAGGATATCCATCAGAAATTAATTTAGTTAGCAATATGAGTTTACCTTGAAGAG 1079
 Db 1273 ACAGGAGACCATCTGAAAGTCGATCGGTTTAAAGACGACAAATTAACCTGCGAGGT 1332
 Qy 1080 AAGGATTAAGTTCAATATCTGATATGACGACCTAATATTATTCGACGTTAAAC 1139
 Db 1333 AAGACGATGTTTACA-----ATACCAAGTCCAAATATATACCTGCTCCTGTTGAT 1386
 Qy 1140 AATGAGAAACGCTATATTTATGATCAAAAACAAGATCAATATCTTCGATGAC 1199
 Db 1387 AACGGTGAACCTGATTTTGGAGATCAAAAAACGACGTTTACCTTATCAACAC 1446
 Qy 1200 ATTACCAAGGCGGAGTGTCTTATTTGAGGTAATTTTACGATATCCAAATCT 1259
 Db 1447 ATCAACCAAGGTGCGGCGGTTGATTTGAGGTAATTTTACGATATCCAAATCT 1506
 Qy 1260 AACCAACTTGGCAAGAGCTGGACATACATGTAAGTAAGTAAGCACTTACGAA 1319
 Db 1507 AATGACATTTGGCAAGGTCCTGAGTGAATGTAAGTAAGTAAGCACTTACGAA 1566
 Qy 1320 GTAATGCGTGAACATATGATGATCTTCTAATAATTTGTAAGAACATTCGACTCA 1379
 Db 1567 GTAATGCGTGAACATATGATGATCTTCTAATAATTTGTAAGAACATTCGACTCA 1626
 Qy 1380 GGCATAAGGGAATAAAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1439
 Db 1627 GGCATAAGGGAATAAAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1686
 Qy 1440 CAGGACGACATCAAGGCAACAAAGCCTTTAGTAATGCTTGGTTAGCGGACA 1499
 Db 1687 CAGGACGACATCAAGGCAACAAAGCCTTTAGTAATGCTTGGTTAGCGGACA 1746
 Qy 1500 GGGACCTTCAATTAACAATGATTAACAATGATTAACAATGATTAACAATGATTAACA 1559
 Db 1747 GGTACGCTTCAATTAACAATGATTAACAATGATTAACAATGATTAACAATGATTAACA 1806
 Qy 1560 CGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1619
 Db 1807 CGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1866
 Qy 1620 GACGAGGGGACATGATGTAACATATACATCACTCAAGCGCTAATGCTATTAAT 1679
 Db 1867 GATGAGGGGACATGATGTAACATATACATCACTCAAGCGCTAATGCTATTAAT 1926
 Qy 1680 GGGACGAAAGCATTTGTTCACTTAATGAAATTAATTAATTAATTAATTAATTAATTA 1739
 Db 1927 GGGACGAAAGCATTTGTTCACTTAATGAAATTAATTAATTAATTAATTAATTAATTA 1977
 Qy 1740 GAAATTCGCTCAACGCGTGTGTTGGGGAACAGTAATAATTAACCATGCGGATTA 1799
 Db 1978 ---ATTGCTTTAATGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGG 2034
 Qy 1800 AACCTTATTTAACAACAACAACAAGATGTAATTTGCTACTTCAAGTGGTCAAT 1859
 Db 2035 AATGTACTTCAATCGCTTAACAACAAGATTAATTAATTAATTAATTAATTAATTAATTA 2094
 Qy 1860 TTAAGAGCGATTAATACCAACAAGGTAATTAATTTTACGCGGTAGACGACACCG 1919
 Db 2095 TTAAGAGCGATTAATTAACCAAGCGGTAGACGCGGTAGACGCGGTAGACGCGGTAG 2154
 Qy 1920 CAGCGCTCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1979
 Db 2155 CAGCGCTCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2208
 Qy 1980 AATGAGGAGTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2039
 Db 2209 GTGGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2268
 Qy 2040 GGGGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2099
 Db 2269 GGGGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2328
 Qy 2100 AATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2159
 Db 2329 AATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2388
 Qy 2160 TCAGATTTGACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2219
 Db 2389 TCAGATTTGACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2448
 Qy 2220 AATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2279
 Db 2449 AATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2508
 Qy 2280 AATGTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 2339
 Db 2509 AATGTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 2568
 Qy 2340 ACATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2399
 Db 2569 ACATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2628
 Qy 2400 ACAGTGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2459
 Db 2629 ACAGTGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2688
 Qy 2460 TTAATAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2519
 Db 2689 TTAATAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2748
 Qy 2520 GAAATTCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2579
 Db 2749 GAAATTCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2808
 Qy 2580 AGTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2636
 Db 2809 AGTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2868
 Qy 2637 CGTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2696
 Db 2869 CGTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2928
 Qy 2697 GTAATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2756
 Db 2929 GTAATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2988
 Qy 2757 AAAAGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2816
 Db 2999 AAAAGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3048
 Qy 2817 AACAGCGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2876
 Db 3049 AACAGCGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3108
 Qy 2877 CAACCGTTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2936
 Db 3109 CAACCGTTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3168

QY 2937 TTACGTTAATAATTAGTGAAGATGCGCAATTCGGTTCGATACCCATATAAG 2996
 Db 3169 TGGGTTAATAATTAGTGAAGAAAAAGGGAATTCCTCGCATATCAATAAAG 3228
 QY 2997 CAGGAATTGCAACAATATTAGTAAGAGCAGACAGACAGAAAGCAATTAGAGCCAA 3056
 Db 3229 CAGGAATTGCGCAATATTAGTAAAGCAGACAGACAGAAAGCAATTAGAGCCAA 3288
 QY 3057 CAAGTGAACCGACTGCTTAAACA CAACAGGTGAGCCAAAAGTGCCTCAAGAGACA 3116
 Db 3289 CAGCTGAATGACTACTTAAACAAAACAAAGGCTAAAGGGCTCAAA----- 3342
 QY 3117 GCGAGACAGCGTTTCTGATATCCCTCCCTGATCAAGCTGTTAAAGCATTAAGACC 3176
 Db 3343 ---AGAGCGCGCTTCTGATATCCCTCCCTGATCAAGCTGTTAAAGCATTAAGACC 3399
 QY 3177 AAACAGCTGAATGACTGCTGAACAA-----AAAGTGAAGCAAAACAAAG 3231
 Db 3400 GAATCTGACAGATTAATGCTCCACAGCAGAGTGCACAGCGTGCAAATGCAAGTA 3459
 QY 3232 TCGGCTCAAAAGACA-----GTGTTTCTGATCCCTGCTGATCAAGCGCTTCGCA 3287
 Db 3460 ACTGACTTAACCAAAAGACAGCAAGTTAAACCACTCAAGTAAAGCAATTTAGTC 3519
 QY 3288 TTAGAAGCCGCACTTGAGGTTATGATGCCCCAC----- 3321
 Db 3520 TTGGCAACTGCAATGGTGAAGAAACCGCTCAGATTGATTGCTAATCAAAATTA 3579
 QY 3322 -----AGCAATCGGA 3332
 Db 3580 GCTCAGTTGAATTTAGCAACAACATGAAAAAGCTTGAAGTGCAGAGCAGAA 3639
 QY 3333 AAAGATGCTAGCTCAAGAAAGAAAGCAACCAACCAAAAGCACTTGAAGC 3392
 Db 3640 AAAGAGGTAAAGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGC 3699
 QY 3393 CGTATTCAAATAGTGCCTATGCAATTTCTGCAACAGTAATAGTATGCTTCTGTT 3452
 Db 3700 CGTATTCAAATAGTGCCTATGCAATTTCTGCAACAGTAATAGTATGCTTCTGTT 3759
 QY 3453 CAAGTGAATTTAGATGCTTTTGTGATCAAGCAATTCGCGGTGCAAAATATC 3512
 Db 3760 CAAGTGAATTTAGATGCTTTTGTGATCAAGCAATTCGCGGTGCAAAATATC 3819
 QY 3513 GCACAGATTAAGAGCTATGATTCGATTCGCTTCGCTTATCAGCAGCAAGAACG 3572
 Db 3820 TCACAGATTAAGAGCTATGATTCGATTCGCTTCGCTTATCAGCAGCAAGAACG 3876
 QY 3573 AACTTACGTCAATTTGCGGTGCAAAAGCTTACCTAATGAGCAATTTGCGGTTC 3632
 Db 3877 AACTTACGTCAATTTGCGGTGCAAAAGCTTACCTAATGAGCAATTTGCGGTTC 3936
 QY 3633 TCGCATGCGCTTCAGATTAATCCTTGTGATGAACAGTTAAATCAAGCGCAATTAAGC 3692
 Db 3937 TCGCATGCGCTTCAGATTAATCCTTGTGATGAACAGTTAAATCAAGCGCAATTAAGC 3996
 QY 3693 ATGATGTCGCGTTTTCGCAATATCAATGCGGCGATTTACATTTGCTTAAAGTGGGA 4056
 Db 3753 ACGGGAATCAGTGCAGTAATGCGTGAAGAAAGCAAGCGCAAAATTTCAATGAAAGCG 3812
 Db 4057 ACGGGAATTTAGTGCAGTAATGCGTGAAGAAAGCAAGCGCAAAATTTCAATGAAAGCG 4116
 QY 3813 ATAAATTAATGCGTAATGCAAGTTATCAGTTCCGTTAAGGCAATGCGCATTCAGCT 3872
 Db 4117 ATAAATTAATGCGTAATGCAAGTTATCAGTTCCGTTAAGGCAATGCGCATTCAGCT 4176
 QY 3873 TATTTGAGGTTAATGCTATTTATTTATGAAAGTGAATTTATCAATCTGAGGAAGTGA 3932
 Db 4177 TATTTGAGGTTAATGCTATTTATTTATGAAAGTGAATTTATCAATCTGAGGAAGTGA 4236
 QY 3933 GTGAAAGCGCTAGCTTGAATTTATTCGCTAATATGCTGCAATGCAATGATTAACA 3992

Db 4237 GTGCAACACGAGCTTGCATTTAATGCTAATATGCTGAGTACGGGTGATTAAGC 4296
 QY 3993 TTACTGCGACGAAATATCAGCGTTAAGCTTATTCCTGCTAATTAAGTGAATGTT 4052
 Db 4297 TTATCCCGACGAGAAATATGCGCTTAAGCTTATTTCTGCTAATTAAGTGAATGTT 4356
 QY 4053 TCAGGCTAAGTGAACCAAGGTAATCTCAGCGTGTGCAACACATTTGAGCGT 4112
 Db 4357 TCAGGCTAAGTGAACCAAGGTAATCTCAGCGTGTGCAACACATTTGAGCGT 4416
 QY 4113 TATGCAAAAGAGTGGGATTAAGGCAAAATTTACATTTCCAAATTCGCTTTT 4172
 Db 4417 TATGCAAAAGAGTGGGATTAAGGCAAAATTTACATTTCCAAATTTCTGCTTTT 4476
 QY 4173 ATCTCAAAATCTCAAGGTTCAACTTCGCAACAGCAAAATGCGGTGAATTCGCG 4232
 Db 4477 ATCTCAAAATCTCAAGGTTCAACTTCGCAACAGCAAAATGCGGTGAATTCGCG 4536
 QY 4233 TATGTTGGTAAATCAATATTTAT 4262
 Db 4537 TATGTTGGTAAATCAATATTTAT 4566

Search completed: June 24, 2004, 19:40:30
 Job time : 13149 secs

Db 5799 GGGAGGTGGCGCAAAATGACAAAGCTCCGATGATTTGGATCTGTCT 5740
 Qy 300 CGTAACGGCGGAGACCTTGGTTAAATCAATATATGAGAGGTGGCAATACGTA 359
 Db 5735 CGTATGCTATGCTGATTAAGCGATCAATATATGAGTGTGCTCAATGGG 5680
 Qy 360 GGAATVACAGATGTTATTTGGTCAGAGGGAACACCCGATCAACCTGTCTTACT 419
 Db 5679 GGAATCACTATGCTATTTGGTCAGAGGCAAAATCCGATCAACCTGTCTTACT 5620
 Qy 420 TTTAGATGTTAAAGAAATTAATCAAAAAAATTAATTAATCTTTGAGAGCAT 479
 Db 5619 TATCAATATGTTAAAGTATATATTAACCTGATCAAGCAATCTATATATGAGGAT 5560
 Qy 480 TACCAATCCAGATTAATCAATATGTTAGAGAGGCGCTCAATTTGATGATCTTGG 539
 Db 5559 TACCAATGCTGCTTATCAATATTTGTAATGTAAGTCTCTTTGATATGACAGA 5500
 Qy 540 AATATGAATGCACTTATCAATATGAAACAAATATCCAGAACGTGTTGTAATCCG 599
 Db 5499 AGTATGAAGGAGACAGCTATATCAGAACAAAGAAATACCCGATGGTGGGATGGT 5440
 Qy 600 TCTGACAGGAGTTTGGCGAAATGATCAAGAAAGCGACAAAGTGGCGGTGAT 659
 Db 5439 TCTGATGATTAATTTGGAATGAAACAAAGATTAAGAACTCAAGTATGATCACTTAT 5380
 Qy 660 CATTATCTGACAGCTGCAATACACACATCAAGTGGAGAGGTTATGATTCGAT 719
 Db 5379 GATATCTTATGCGGGGAATCCCATATGCAAACTGTTGAGGAAATGATGATTTGAT 5320
 Qy 720 TTGGAGAGGCAATTTGTAACCGGAGAAATATGTCATTAACGATGAGGCTCAAG 779
 Db 5319 TTAAGTGGCATATGATTAAGAAATGATTAAGACCTATGACCAAGAGATCATTT 5260
 Qy 780 GGGAGAGTGTCTCCGATGTTATTAATGATGCTGAAAAAATATGATTAAT 839
 Db 5259 GAGACAGTGGTGCATATGTTATTAATGATGCTGACCGGAAAAATGCTTATTAAT 5200
 Qy 840 GGGATTTACGAGGAGGCAACCTTTGAGAGGAAAGAAATGGGTTTCAATGGTGGC 899
 Db 5199 GGTGATTTGGGTAACGGAACCCATCTTGGTGAAGAGATATTTTCACTGATCTGT 5140
 Qy 900 AATCTTATTTTGAATTTTGAAGAGAT-TTAATATCATCACTTATTAACCCGAGC 958
 Db 5139 AATATTTATTTGATTAATTTTCTGCTAATCTTAAGACATATGATTTGCA 5080
 Qy 959 TGGTATGAGGTATACAAATTAAGTGA-----ATGATATGTCAGGGGTCTATA 1010
 Db 5079 ACTCCATAAATCAAAATCAAGTCACTACATTTCAATCAAAAATGATGGGTCAAGC 5020
 Qy 1011 ACTCAGAAATCAAGAAATCAATCAAGAAATTAATTAATGATTAATGATTAATCT 1070
 Db 5019 ACTATATGATGAGAGTGAATTAATTAAGTCAAACTGATTAATTAATTAATCT 4960
 Qy 1071 TTGAAGAGAGATTAAGTCAATATCTGATATGACGAGCTTAATTAATTAATCTGCA 1130
 Db 4959 ATGAAAGGACGAGAGCTGTCTTACT-----AATGGTGTGAACAAATTAATCTCC 4906
 Qy 1131 GGTAAACAAATGAGAAACGCTATATTTATGATCAAAATCAAGAGATCAATATCTTC 1190
 Db 4905 GATTAATGTCAGGCAAAACCTTATTTATGATTAAGAAAGGGGTGCTAATCTTT 4846
 Qy 1191 GATCTGATCAATTAACAAAGGCGGATGCTTATTTTGAAGGATTAATTTAATCAATCT 1250
 Db 4845 AGTATATATATGATCAAGGTGCAAGGGCTTATTTTGAAGGAGATTTTATATGTTG 4786
 Qy 1251 CCAATTTCTAACAACTTTGCAAGAGGCTGGGATCAATGATTAATTAATTAATGACGCT 1310
 Db 4785 CCAATCCATATGAACCTGGAAGGGCGGGATATCTAATCAATGATTAAGAGCCCTTA 4726
 Qy 1311 ACTTGAAGTAAATGAGTGGAAATGATCACTTTCAAAATTTGGTAAAGGACATG 1370
 Db 4725 ATTGGAATGATGAGGTGAAATGATGCTTTCTAAATTTGGTAAAGGACCTTA 4666
 Qy 1371 CAGCTTCAAGCCAAAGGGGAAATTAAGGTGATCAAGGCTAGGAGGATTAATCT 1430
 Db 4665 CATTTTCAAGCAAAAGGTAATTAATTAAGGTGATCAAGGCTAGGAGGATTAATCT 4606
 Qy 1431 TTGAGACAGAGGAGAGATCAAGGCAAAACAAAGCTTTAGTAAATTTGGTGT 1490
 Db 4605 TTGAACAAAGAGAGAGATCAAGGCAAAACAAAGCTTTAGTAAATTTGGTGT 4546
 Qy 1491 AGGCGAGAGGAGCTGTCAATTAAGATGATTAACAAATTTGATCCGATTAATTTAT 1550
 Db 4545 AGGCGAGAGGAGCTGTCAATTAAGATGATTAACAAATTTGATCCGATTAATTTAT 4486
 Qy 1551 TTGGCTTGTGGGTGCTTATGATCTTAAGAGGAGCTCAATTAACCTTTAAGATATC 1610
 Db 4485 TTGGCTTGTGGGTGCTTATGATCTTAAGAGGAGCTCAATTAACCTTTAAGATATC 4426
 Qy 1611 CAAATATCGAGAGAGGAGATGATGATGATCAATTAACCTCAAGCGCTAATGTC 1670
 Db 4425 CAAATATCGAGAGAGGAGATGATGATGATCAATTAACCTCAAGCGCTAATGTC 4366
 Qy 1671 ACTATTAATCGAGAGAGATGATGATGATCAATTAACCTCAAGCGCTAATGTC 1730
 Db 4365 ACTATTAATCGAGAGAGATGATGATGATCAATTAACCTCAAGCGCTAATGTC 4306
 Qy 1731 TACAGAAAGAAATTTGCTCAAGCGCTTGTGGGCAACAGATTAATTAACCTAAT 1790
 Db 4305 TACAGAAAGAAATTTGCTCAAGCGCTTGTGGGCAACAGATTAATTAACCTAAT 4246
 Qy 1791 GGGCGATTAACCTTATTTAATTAACCAACAGAGATGATTAACCTTATGAT 1850
 Db 4245 GAGAGATTAACCTTATTTAATTAACCAACAGAGATGATTAACCTTATGAT 4186
 Qy 1851 GATTAATTAATTAAGGAGATTAATTAACCAACAGAGATGATTAACCTTATGAT 1910
 Db 4185 GAGAGATTAATTAAGGAGATTAATTAACCAACAGAGATGATTAACCTTATGAT 4126
 Qy 1911 CCGACAGCGAGCGCTCAATCAATTAATTAACCTTATGATTAACCTTATGAT 1970
 Db 4125 CCGACAGCGAGCGCTCAATCAATTAATTAACCTTATGATTAACCTTATGAT 4072
 Qy 1971 CAGGCGAAATTTGTGGATCAAGATGATCAACCTTAATTAACCTTATGAT 2030
 Db 4071 CAGGCGAAATTTGTGGATCAAGATGATCAACCTTAATTAACCTTATGAT 4012
 Qy 2031 CAAATTAAGGAGAGATGAGGCTTGTGGATCAAGATGATCAACCTTAATTAACCT 2090
 Db 4012 CAAATTAAGGAGAGATGAGGCTTGTGGATCAAGATGATCAACCTTAATTAACCT 3952
 Qy 2091 ACAGTCAAGATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2150
 Db 3951 ACAGTCAAGATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3892
 Qy 2151 TGAAGCGTCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2210
 Db 3891 TGAAGCGTCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3832
 Qy 2211 AAGATTAATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2270
 Db 3831 AAGATTAATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3772
 Qy 2271 GCAAGCGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2330
 Db 3771 GCAAGCGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3712
 Qy 2331 AGCAATTAATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2390
 Db 3711 AGCAATTAATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3652
 Qy 2391 TCAAGTCAAGGAGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2450
 Db 3651 TCAAGTCAAGGAGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 3592

QY	1851	GGCTCAATTTAAAGGGGAGATTACCGCAAAACAAAGGTAACTATTTTTCACGGGTGA	1910
Db	210661	GGCCAAACCTTAAAGGCGATTACTCTCAAGAGGCGGCACTTATGCTTTAGCGTGC	2106022
QY	1911	CCGACACCGCAGCGCTTACAAATCATTTTAAATTAAGTTGCTCGAAGATGSAAGGTATACCA	1970
Db	210601	CCAAAGCCACAGGCTTACAAATCATTTAA-----TGGCCAAAGAGCTTTGGCGACT	2105488
QY	1971	CAAGCGAAATGTGTGGGATCAGATTGSAATCAACGTACATTTAAAGCTGAAATCTTC	2030
Db	210547	CAAGCGAAATGCTTATGTATGTAGCGATTGSAATCAACGCACTTTTAAAGCTGAAATCTTC	2104888
QY	2031	CAATTTAAAGGCGAAGTGGCGGTGTCTCGCAATGTTCTTCAATTGAGGGAATTTGG	20930
Db	210487	CAATTTAAAGGCGAAGTGGCGGTGTCTCGCAATGTTCTTCAATTGAGGGAATTTGG	2104288
QY	2091	ACAGTCAGCAATATGCAAAATGCGACATTTGGTGTGTGCGCAATCAACAAATACCAT	2150
Db	210427	ACATTCAGCAATATGCGCAACGCGACATTTGGTGTGTGCGCAATCAACAAATACCAT	2103688
QY	2151	TGCGCGGTTCGATGSAACAGATTAAACACTTGTCAAAAAGTGAATTTAACCGATACA	2210
Db	210367	TGCGCGGTTCGATGSAACAGSCTTAAACGATTTGAAACAGCAATTTAACCGACACA	2103088
QY	2211	AAAGTTATTAATTCATACCAAAAACACAATCAATGSGCTCATTTAATTTACGTATAT	2270
Db	210307	AAAGTTATTAATTCATACCAACAAACAAATCAATGSGCTCATTTAATTTACGTATAT	2102488
QY	2271	GCACCGGCGAATGTTAAAGGTTTAGCAAACTTAATGCGAATGTCACTTAACAATCAC	2330
Db	210247	GCACCGGCGAATGTTAAAGGTTTAGCAAACTTAAATGCGAATGTCACTTAACAATCAC	2101888
QY	2331	AGCCATTTACATTAAGCAACAATGCAACCGCAATAGGCAATATTCGACTTCCGAGAT	2390
Db	210187	AGCCATTTACATTAAGCAACAATGCAACCGCAATAGGCAATATTCGACTTCCGAGAT	2101288
QY	2391	TCAACTGCAACGCTGSAATATGCAACTTGAACGATATGTGCATTTAACGATTCAGCT	2450
Db	210127	TCAACTGCAACGCTGSAATATGCAACTTGAACGATATGTGCATTTAAGATTCAGCT	2100688
QY	2451	CAATTTCTTTAAAAACGCCATTTTTCGACCAATTCAGGSAAGCAAAAGGCAACA	2510
Db	210067	CAATTTCTTTAAAAACGCCATCTTTCGACCAATTCAGGSAAGCAAAAGGCAACA	2100088
QY	2511	GTCGCGTGGAAAAATGCGACTTGGCAATGCTAGGATACATCTTGAGAAATTAAG	2570
Db	210007	GTCGCGTGGAAAAATGCGACTTGGCAATGCTAGGATACATCTTGAGAAATTAACA	2093488
QY	2571	CTAATTAACAGTACGATCAAGTTAAATTCAGCTTATTCAGTACGCTCAAAATAGCGCA	2630
Db	209947	CTAATTAACAGTACGTTACGTTAAATTCAGCTTATTCAGTACGCTCAAAATAGCGCA	2098888
QY	2631	-----GTCGCGCTTCAATTAAGAGCGAAACAACGCGCAATCGGCGAAGCAATGTTTC	2684
Db	209887	CGCGCTGTCGCGCTTCAATTAAGAGCGAAACAACGCGCAATCGGCGAAGCAATGTTTC	2098288
QY	2685	AACACATTCAGTAAATGTTAAATTTGATGAGGGAAGGACATTCCAATTTACTTCATCT	2744
Db	209827	AACACATTCAGTAAATGTTAAATTTGATGAGGGAAGGACATTCCAATTTACTTCATCT	2097688
QY	2745	TTATTTGGCTATTAAGAGCATTAATTTAAATATTCCAATGACGCTGAGGCGCATTTACATA	2804
Db	209767	TTATTTGGCTATTAAGAGCATTAATTTAAATATTCCAATGAGCGCTGAGGCGCATTTACATA	2097088
QY	2805	TTATTCGTTTCGACACAGCGCAAGAACCCGGAACCCCTTAGCAATTAATCTTGGTTGAA	2864
Db	209707	TTATTCGTTTCGACACAGCGCAAGAACCCGGAACCCCTTAGCAATTAATCTTGGTTGAA	2096488
QY	2865	AGCAAGATTAACAACCGTTATCAGTAGGCTCAATTTACTTTAGAAATGACACAGTT	2924
Db	209647	AGCAAGATTAACAACCGTTATCAGTAATCACTCACTTACGTTAGAAATGACACAGTT	2095888

QY	2922	GATCGACGTCGCTTACGCTTAAATATAGGAAAGATAGAGCGAATCCGCTTCGCTAAC	2994
Db	209587	GATCGGCGTCATTACGCTATGATTTAGTGAAGAAATATCGGAATTCGCTTCGCTAAT	2095282
QY	2985	CCAAATAAAAGACAGCAATTTGCAACATATATTAGTAGAGCAGCAAGCAAGCAACA	3044
Db	209527	CCAAATAAAAGACAGCAATTTGCAACATATATTAGTAGAGCAGCAAGCAAGCAACA	209468
QY	3045	TTAAGACCCAAACAGTTGAAACCGACTGCTAAACACAAACAGGTAGCCAAATGTCGCG	3104
Db	209467	TTAAGACCCAAACAGTTGAAACCGACTGCTAAACACAAACAGGTAGCCAAATGTCGCG	209408
QY	3105	TCGAAGAGAGCAGCAGAGCGGCTTCGTATACCCCTGCTGATCGAAAGCGTTAAAC	3154
Db	209407	TTAAAGA-----AGAGCGGCTTTCCTATACCTGCTGCTGAAGCTGTTAAAC	209357
QY	3165	GCATTAGAGCCAAACAGCTGAACCTGCTGTAACCAAAAAAGTAGGCAAAAAACA	3224
Db	209356	GCATTAGAGCCAAACAGCTGAACCGAATCTTAAACCAAAAAAGTAGGCAAAAAACA	209297
QY	3225	AAAAAAGTCGGCTCAAAAAGCA-----GTGTTTCTGATCCCTGCTTATCA	3275
Db	209296	AAAAAAGTCGGCTCAAAAAGCAATGAGAGCGGTGTTTTCGATCCCGCGCTGATCA	209237
QY	3276	AGCGCTG---CGCATAGAGCCGCACTTAGGTTATTTGAGCCCGCAACAGCAATGGGA	3332
Db	209236	AGCGCATGAGCGCTTTTAAAGCCGCACTTAGGTTATTTAGCCCAACCGCAAGTGGGA	209177
QY	3333	AAAGATCGCTAGCTCAAGAGAGAGCGGAAACAAACGAAACAAAAAGACTTATCAGC	3392
Db	209176	AAAGATCGCTCAAGCTCAAGAGAGAGCGGAAACAAACGAAACAAAAAGACTTATCAGC	209126
QY	3393	CGTTATTCAATAGTGGCTTATCAATATATCTGCAACGTAAATAGTATGCTTCTGT	3452
Db	209125	CGTTACTCAAAATAGTGGCTTATCGAGTTCCTGCAACGTAAATAGTATGCTTCTGT	209065
QY	3453	CAAGATGATTTAAACGCTCTTTTGTAGATCAAGCAACATCGCGCTGAGCAAAATAC	3512
Db	209065	CAAGATGATTTAAACGCTCTTTTGTAGATCAAGCAACATCGCGCTGAGCAAAATAC	209006
QY	3513	GCAACAGATTAAGAGCGCTATGATCTGATCGCTCCGCTCTTACAGCAGCAAAAAAG	3572
Db	209005	GCAACAGATTAAGAGCGCTATGATCTGATCGCTCCGCTCTTACAGCAGCAAAAAAG	208943
QY	3573	AACCTACGCAATTTGGGAGTGCAAAAACCTTAGCTAATGACGAATTTGGGAGGTTTTC	3632
Db	208948	AACCTACGCAATTTGGGAGTGCAAAAACCTTAGCTAATGACGAATTTGGGAGGTTTTC	208889
QY	3633	TTCGATAGCGCTTGCAGCAATTAACCTTGAAGAACAGTTAAATTCACCGACATTAACG	3692
Db	208888	TTCGATAGCGCTTGCAGCAATTAACCTTGAAGAACAGTTAAATTCACCGACATTAACG	208829
QY	3693	ATGATGTCGGGTTTTGGCCCAATATCAATGGGGGCGATTTACATTTTGCTGAACCGTGGGA	3752
Db	208828	ATGATGTCGGGTTTTGGCCCAATATCAATGGGGGCGATTTACATTTTGCTGAACCGTGGGA	208769
QY	3753	ACCGGATCATGTCGAGTAAATTTGCTGAGAACAAAGCCGAAAAATTCATCGAAAAACG	3812
Db	208768	ACCGGATCATGTCGAGTAAATTTGCTGAGAACAAAGCCGAAAAATTCATCGAAAAACG	208709
QY	3813	ATAAATTTAGGCGTGAATGCAAGTTATACAGTCCGTTTAGGGCAATTTGGGCACTTACGCT	3872
Db	208708	ATAAATTTAGGCGTGAATGCAAGTTATACAGTCCGTTTAGGGCAATTTGGGCACTTACGCT	208649
QY	3873	TATTTGAGGTAAATGCGTATTTATGCAACGCGGAATATCAATCTGAGGAGCTGAGGA	3932
Db	208648	TATTTGAGGTAAATGCGTATTTATGCAACGCGGAATATCAATCTGAGGAGCTGAGGA	208589
QY	3933	GTGAAAAAGCCTAGCCTTGATTTATGCTATATGCTTGGCAATTCAGTTGATTTACA	3992
Db	208588	GTGAAAAAGCAGAGCTTGATTTATGCTATATGCTTGGCAATTCAGTTGATTTACG	208529
QY	3993	TTTATCTCGACAGTAATATACGCTTAAAGCTTATTTCTGCAATTAATGTTATGTTT	4052


```

Db 208528 TTACCCCGACATATATCAGCGTTAGCTTATTTCTGTAATATATGATATCT 208469
Qy 4053 TCACAGCTACGACACACACGCTTAATCTCAGCTTGTGCAACACATTTGGACGT 4112
Db 208468 TCACAGCTACGACACACACGCTTAATCTCAGCTTGTGCAACATCATTTGGACGT 208409
Qy 4113 TATGGCAAAAAGAGTGGGATTTAAGCAGAAATTTTACATTTCCAAATTTCCGCTTT 4172
Db 208408 TATGGCAAAAAGAGTGGGATTTAAGCAGAAATTTTACATTTCCAAATTTCCGCTTT 208349
Qy 4173 ATCTCAAAATCTCAGCTTGTGCAACATCTGCGCAACACCAAAATGTGGCGTGAATTTGGGC 4232
Db 208348 ATCTCAAAATCTCAGCTTGTGCAACATCTGCGCAACACCAAAATGTGGCGTGAATTTGGGC 208289
Qy 4233 TATGGTGTGTAATAATCAATATTTATGCTTATGATTAACAAGTGGCTCAATTC 4292
Db 208288 TATGGTGTGTAATAATCAATATTTATGCTTATGATTAACAAGTGGCGTGAATTC 208229
Qy 4293 AGATCCACCTTTTATTTTCAATAT 4319
Db 208228 AATCCACCTTTTATTTTCAATAT 208202

```

```

RESULT 3
PCT-US04-06615-5
; Sequence 5, Application PC/TUS0406615
; GENERAL INFORMATION:
; APPLICANT: New England Medical Center Hospitals, Inc.
; APPLICANT: Plant, Andrew G
; APPLICANT: Qiu, Jiazhou
; TITLE OF INVENTION: Treatment of IGA1 Deposition Diseases
; FILE REFERENCE: 28154/2268
; CURRENT APPLICATION NUMBER: PCT/US04/06615
; PRIOR FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US 60/453055
; PRIOR FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; LENGTH: 5085
; SEQ ID NO 5
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
PCT-US04-06615-5

```

```

Query Match 4.7%; Score 202; DB 1; Length 5085;
Best Local Similarity 60.9%; Pred. No. 6,1e-40;
Matches 347; Conservative 0; Mismatches 220; Indels 3; Gaps 2;

```

```

Qy 1176 GGATCATTATATCTTCGCACTGACATTACCAAGGCGGCGGTGCTTTATTTGAGGGT 1235
Db 1183 GGAACGCTTACTTAATATATATATCATGATCAAGGCGGCGGCTTTGTTCTTAAAGA 1242
Qy 1236 AATTTACGATATCTTCAATTC---TAAACAACTTGGCAAGAGTGGCATATATGTA 1292
Db 1243 GATTTATGATTAAGGACCTTCTGATGATGACACTGAGAAAGAGTGGCGTTTCGTT 1302
Qy 1293 AGTAAATATGACCGCTTACTTGTGAAGTAAATGCGTGAACATGATGACCTTTCAA 1352
Db 1303 GGTATGAAAAACGTAACGTGAAAGTACATTAACCGGAATGTATGCTTAACTAA 1362
Qy 1353 ATTTGTAAGGAAATGCGTTCACAGCCAAAGGCGAAATTAAGGTTGATACGCGTA 1412
Db 1363 ATTCGCAAGGAAATATATGTAAGGAAAGGGAATTAAGGTTGCTTAAAGTG 1422
Qy 1413 GAGCATGTTAAGTCAATTTTGAAGCAGCGGCAAGCATCAAGGCAACAAAGACCTTT 1472
Db 1423 GGCATGTTAAGTCAATTTTGAAGCAGCGGCAAGCATCAATTAAGTCAAGACCTTT 1482
Qy 1473 AGTAAATGCGCTTGTGTAAGGCGGCAAGGATGTTCAATTAACGATTAACAAATTT 1532
Db 1483 TCACAGTATGATATGTAAGTGTGCTCAAGCTGTTGTAATTAATGATTAAGCAAGTA 1542

```

```

Qy 1533 GATACGATTAATTTTATTTGCGCTTGTGGTGGTGTGATCTTAATCAAGGCGATTC 1592
Db 1543 GATCAAAATTTCAATTTATTTGCTTTAGAGTGGTGGATTAATCAAGGCGATTAAT 1602
Qy 1593 TTAACCTTTAAGTATTCACAAATTAAGGACGAGGCGGCAATGATGTGAACCATATCA 1652
Db 1603 CTCACTTTGAACATTCGTAATTTGATGATGATGATGATGATGATGATGATGATGAT 1662
Qy 1653 ACTGAGCGGTATATGATCAATTTACTTGGGACGAAAGCATGCTTACATTAAGGAAAT 1712
Db 1663 AGCAAACTCTTACGTAATTAATTTACTTGGGAAAGCTTAATTAAGATTCGAATTAAT 1722
Qy 1713 AATATTAATTAATCTGATTAAGAAAGAA 1742
Db 1723 ACTGATTAATTAAGGACGACGATGAA 1752

```

```

RESULT 4
PCT-US04-07001-563/c
; Sequence 563, Application PC/TUS0407001
; GENERAL INFORMATION:
; APPLICANT: BAKARSTZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
; FILE REFERENCE: 28335/38815A
; CURRENT APPLICATION NUMBER: PCT/US04/07001
; PRIOR FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 563
; LENGTH: 14710
; TYPE: DNA
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8250)..(8250)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8293)..(8293)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8350)..(8350)
; OTHER INFORMATION: n = a, c, g, or t
PCT-US04-07001-563

```

```

Query Match 4.1%; Score 177.6; DB 1; Length 14710;
Best Local Similarity 59.3%; Pred. No. 1.1e-33;
Matches 341; Conservative 0; Mismatches 225; Indels 9; Gaps 2;

```

```

Qy 1154 ATATTTTATGATCAAAAACAGATCAATATCTTGCATCTGACATTAAACCAAGGCG 1213
Db 8261 AATATGTTAGTAAAGAAAGTGAACCTTAAATTAATATATATGATCAAGTGC 8202
Qy 1214 GGGTGGCTTTATTTTGAAGGTAATTTTACAGTATCTCAAAATTC---TAAACAACTTG 1270
Db 8201 AGCGGATTTGTTTGAAGGATTAATGAAGTAAAGGACCTTAAAGAAATTAAGCTTG 8142
Qy 1271 GCAAGGCTGGCATATCATATTAAGTAAATTAAGCAAGGCTTACTTGAAGTAAATGCGCT 1330
Db 8141 GAAGGCGCTGGTATATCTTGTGAGGAAACAGTAAAGTGAAGTGAAGTGAAGTGAAG 8082
Qy 1331 GGAACATATGACCTTCTTAATTTGTAAGGAAACATTTGACCTTCAAGCAAGGCGA 1390
Db 8081 TCAATTTGCTTATGATTAATTTGTAAGGAAACATTTGTAAGGAAAGGCGA 8022
Qy 1391 AATTAAGGTTGATTCAGCTTGAAGGATGTTAAAGTCAATTTTGAAGGCGAGGCGA 1450
Db 8021 TATTAAGGTTGCTTAAAGTGGGAGTGGCAAGTGTGTTTAAAGCAACAAACACTAC 7962
Qy 1451 TCAGGCAACAAAGACCTTTAGTGAATTTGCTTGTGTAAGGCGAGGAGGCTGTTCA 1510

```

Db 7961 AGAGC-----AAACGCTTTTGGTCTGTGAGAGAAATGTGAAGGTGCTCTCAACTGTGTCT 7908

Qy 1511 ATTTAAAGATGATTAACCAATTGATATACCAATTAATTTATTTGGGCTTCTGTGGTGTGTG 1570

Db 7907 ACTTAAATGATGATTAACCAAGTAGTACCAATTCATTTACTCTTGGTTTAAAGGCGGTGTG 7848

Qy 1571 CTTAGATCTTAAAGGGCATTTCATTAACTTTAAACGATCCAAATATAGGACGAGGCGAGC 1630

Db 7847 ATTGAGATGCCAATGGCAATATATCTCATCTTTTAAACATATCCGTAATATATGATGATGTCG 7788

Qy 1631 AATGATGTGTAACCAATTAATACACTCAAGCCGCTATATGTCACTAATTAATCTGGGAACGAAG 1690

Db 7787 AAGAGCTAGTAAATCAATATATGACTAATAGCTCAAAATTAATTAACGATTTACGAGAGAGATT 7728

Qy 1691 CATGTGTTCTACCTAATGGAATAATATTAATTAAC 1725

Db 7727 AATTACAATTCACAGTACAGGTCACTATATATACG 7693

```

RESULT 5
PCT-US04-07001-762
: Sequence 762, Application PC/TUS0407001
: GENERAL INFORMATION:
: APPLICANT: BARALETZ et al.
: TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
: FILE REFERENCE: 28335/38815A
: CURRENT APPLICATION NUMBER: PCT/US04/07001
: CURRENT FILING DATE: 2004-03-05
: PRIOR APPLICATION NUMBER: US 60/453,134
: PRIOR FILING DATE: 2003-03-06
: NUMBER OF SEQ ID NOS: 771
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 762
: LENGTH: 5385
: TYPE: DNA
: ORGANISM: H. influenzae
PCT-US04-07001-762

```

Query Match	Score	DB 1	Length
4.18;	177;	DB 1;	5385;

Best Local Similarity 59.3%; Pred. NO. 1.1e-33;
Matches 341; Conservative 0; Mismatches 225; Indels 9; Gaps 2

```

QY 1154 ATATTATGAGATCAAAAACAAGATCATTAATCTTCGGATCTGACATTAACAAGGGG 121
Db 1248 AATGTGTTAGCTTTGAGGAAGTGGAACCTTACCTTAATATATATATGATCAAGGTG 1307
QY 1214 GGGTGTCTTTATTTTAGGGGTATTTTACAGATCTCCAAATTC---TAACAAACTTG 1270
Db 1308 AGCGGATGTTCTTTTGAAAGTGATTTATGAAGTTAAAGGCCTTCAGAAAATAGCACTTG 1367
QY 1271 GCAGAGAGCTGGCATACATGTAGTGAATAGCAACCGTTACTTGGAAAGTAAATGCGCT 1333
Db 1368 GAAGGGGCTGGTATATCTGTTGCGAAGAAAAAAGATTAAGTGAAAGTGATTAACCC 1422
QY 1331 GGAACATGATGCATTTCTPAAATTTGGTAAAGACATTCAGACTTCAAGCCAAAGGGGA 1390
Db 1428 TCAATTTGATGGTTTACCTAAATTTGTTAAAGGAAGCAATTTGTGAAGACAGAGGA 1487
QY 1391 AATTAAGGTTCGATCGAGGTAGGGAAGGTAAAGCATTTTGGACACAGAGCAACGA 1450
Db 1488 TATTAAGGTTCGCTAAAGATGGGGAGAGCAAGTTGTTTAAACAACAACACTAC 1547
QY 1451 TCAGGCAACAACAACACCTTTAGTGAATTTGGCTGGTTAGCGGACAGAGGACGTGTA 1517
Db 1548 AGGAC-----AAGACCTTTTGGCTTCGTAGGAATTTGATAGTGTGTGCTCACTGTTG 1601
QY 1511 ATTAAAGATATAAACAATTGATACGATTAATTTATTCGCGCTTCGTGTGTGTG 1570
Db 1602 ACTTATATGATATACCAAGATGATCCAAATTCATTTACCTTGTTTAGAGGGGTG 1661
QY 1571 CTTAAGTCTTAAGGGCATTCATTAACCTTTAAAGTATCCAAATTAAGCAAGGGGCG 1633

```

Db 1662 ATTAGATGGCAATGGCAATATATCATCTTTGGAACATATCCGAATATATGATATGATGTC 1721

Gy 1631 AATATATGGAACCCATATACATCAATCAAGCCGATAGGACATATATCTGGAAAG 1630

Dd 1722 AAGACTATGATATCTATATATACCTAAGCCCTCAATATATGATTAATGAGACGAGTT 1781

Gy 1631 CATTTCTTACTTAAATGAAATATATTTATTAAC 1725

Dd 1782 AATTACAAATCCAGTCAGGTCCTATATATATAC 1816

```

RESULT 6
PCT-US04-07001-683/c
; Sequence 683, Application PC/TUS0407001
; GENERAL INFORMATION:
; APPLICANT: BAYKATZ et al.
; TITLE OP INVENTION: GENES OF AN OTTIS MEDIA ISOLATE OF MONTYPEABLE H. INFLUENZA
; FILE REFERENCE: 2835/38615A
; CURRENT APPLICATION NUMBER: PCT/US04/07001
; PRIOR FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 683
; LENGTH: 191996
; TYPE: DNA
; ORGANISM: H. influenzae
PCT-US04-07001-683

```

Query Match	4.1%; Score 177; DB 1; Length 191996
-------------	--------------------------------------

Best Local Similarity 59.38; PRed.NO. 4.3e-33;
Matches 341; Conservative 0; Mismatches 225; Indels 9; Gaps 2

1154 ATATTTATGATCAAAACAGGATCATTAATCTTGGCTGTGACATTAAACCAAGGGC 1213

Db 170400 AAATGTTACGTTGAAGGAAGTGAAACCTTACCTTAAATTAATATATCGATCAAGTGC 170341

1214 GGGTGGCTTTATTTTGAGGGAATTTTACAGTATCTCCAATTC--TAACCAACCTTG 1270

D5 170340 AGCGGATTGTTCTTGAAGTGATATGAAGTTAAAGCACTTCAGAAATACGACTTG 170281

1271 GCAAGAGCTGGCATACATGTAGTGAAATAGCACCGTACTTGGAAATAATGGCGT 1330

Db 170280 GAAAGGGCTGTATATCTGTCTGAAGGAAACAGTAAAGTGAAGTCATAACCC 170221

1331 GGACATGATCGACTTCTTAAATTGCTAAAGGACATTGCACGTTCAAGCCAAAGGGA 1330

Db 170220 TCAATTGATCGTTAGCTBAATTGCTAAAGGAAGCTAATTGTTGAAGGAGAGGAGA 170161

1391 AAATTAAGCTTCGATCAGCGTAGCGCATGGTAAGTCATTTTCGAGCAGCAGGCAGACGA 1450

Db 170160 TATAAGGTCGCTAAAAGTGGCCATGGCCACTGTGTTTAAACACACAACTAC 170101

1451 TCACGCCAACAAACAGCCTTAGTGAATTGGCTTGGTTACCGGCAGAGGACTGTCA 1510

Db 170100 AGGAC-----AACACGCTTTCCTCTGTAGGAATGTAACTGTCGCTCACTGTTGT 170047

1511 ATTTAAACGATGATTAACAATTGATGACCGATTAATTATTATTTTCGCGCTTCGTGCTGCTG 1520

170646 ACTTATGATGATACCAAGTAGATCCAAATTCATATTAATTGCTTTAGAGAGCGCTCG 16998T

1571 1630

[illegible]

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

1000

1. The first step is to identify the problem or question that needs to be addressed. This involves understanding the context and the specific requirements of the task.

RESULT 7

US-10-215-432A-29

; Sequence 29, Application US/10215432A

; GENERAL INFORMATION:

; APPLICANT: KMEC, ERIC B.

; APPLICANT: PAREKH-OLMEDO, HETAL

; TITLE OF INVENTION: COMPOSITION AND METHODS FOR THE PREVENTION AND TREATMENT OF

; TITLE OF INVENTION: HUNTINGTON'S DISEASE

; FILE REFERENCE: NaPro-10 US

; CURRENT APPLICATION NUMBER: US/10/215,432A

; CURRENT FILING DATE: 2002-08-07

; PRIOR APPLICATION NUMBER: US 60/310,757

; PRIOR FILING DATE: 2001-08-07

; PRIOR APPLICATION NUMBER: US 60/310,889

; PRIOR FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: US 60/310,770

; PRIOR FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: US 60/337,219

; PRIOR FILING DATE: 2001-12-04

; NUMBER OF SEQ ID NOS: 63

; SOFTWARE: PatentIn ver. 3.2

; SEQ ID NO 29

; LENGTH: 561

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: Repeat region

; LOCATION: (51) (411)

; OTHER INFORMATION: this region may encompass 20-120 car repeats according to

; OTHER INFORMATION: the specification as filed; see specification for detailed

; OTHER INFORMATION: description of specific embodiments

US-10-215-432A-29

Query Match

1.1%; Score 48.6; DB 7; Length 561;

Best Local Similarity 34.0%; Pred. No. 0.047;

Matches 101; Conservative 63; Mismatches 132; Indels 1; Gaps 1;

QY 2953 TGAAGATGATGCGGATTCGCTTGCCTAACCATAAAGAGAGAGATTCACATG 3012

DB 20 TATGAGGCGCTTGAGTCCCTCAAGTCCTCCARCARCARCARCARCARCARC 79

QY 3013 ATTGTAG 3072

DB 80 A-RCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARC 138

QY 3073 CTAAACACAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3132

DB 139 CARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARC 198

QY 3133 CTGATACCTGCTGATCAAGCTGTAAAGCATTAGAGAGAGAGAGAGAGAGAG 3192

DB 199 CARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARC 258

QY 3193 CTGCTGAACCAAAAAGTAAGCAAAAAGTAAGCAAAAAGTAAGCAAAAAGTA 3249

DB 259 CARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARC 315

RESULT 8

US-10-859-198-1510

; Sequence 1510, Application US/10859198

; GENERAL INFORMATION:

; APPLICANT: Wyeth

; APPLICANT: Mounts, William

; APPLICANT: Murphy, Ellen

; APPLICANT: Whitley, Maryann

; TITLE OF INVENTION: Nucleic Acid Arrays for Detecting Multiple Strains of a Non-Viral

; FILE REFERENCE: 031896-014000 (AM101085)

; CURRENT APPLICATION NUMBER: US/10/859,198

; CURRENT FILING DATE: 2004-06-03

Query Match

1.1%; Score 48.6; DB 7; Length 561;

Best Local Similarity 34.0%; Pred. No. 0.047;

Matches 101; Conservative 63; Mismatches 132; Indels 1; Gaps 1;

QY 2953 TGAAGATGATGCGGATTCGCTTGCCTAACCATAAAGAGAGAGATTCACATG 3012

DB 20 TATGAGGCGCTTGAGTCCCTCAAGTCCTCCARCARCARCARCARCARCARC 79

QY 3013 ATTGTAG 3072

DB 80 A-RCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARC 138

QY 3073 CTAAACACAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3132

DB 139 CARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARC 198

QY 3133 CTGATACCTGCTGATCAAGCTGTAAAGCATTAGAGAGAGAGAGAGAGAGAG 3192

DB 199 CARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARC 258

QY 3193 CTGCTGAACCAAAAAGTAAGCAAAAAGTAAGCAAAAAGTAAGCAAAAAGTA 3249

DB 259 CARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARCARC 315

; PRIOR APPLICATION NUMBER: US 60/475,871

; PRIOR FILING DATE: 2003-06-05

; NUMBER OF SEQ ID NOS: 282011

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1510

; LENGTH: 4647

; TYPE: DNA

; ORGANISM: Staphylococcus aureus

US-10-859-198-1510

Query Match

1.1%; Score 48.2; DB 7; Length 4647;

Best Local Similarity 44.3%; Pred. No. 0.13;

Matches 197; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 2958 ATAAAGAGAGAGATTCGATCATGATTTAGTAGAGAGAGAGAGAGAGAGATTA 3047

DB 319 ACAAAGAGAGAGATTCGATCATGATTTAGTAGAGAGAGAGAGAGAGAGATTA 378

QY 3048 GAAAG 3107

DB 379 GAAAG 438

QY 3108 AGAAG 3167

DB 439 GAAAG 498

QY 3168 TTAAG 3227

DB 439 GAAAG 558

QY 3228 AAG 3287

DB 559 GAAAG 618

QY 3288 TTAAG 3347

DB 619 ACAAAG 678

QY 3348 CAAG 3407

DB 679 GCTGAG 738

QY 3408 GCGTTAACAAGATTTCTGCAACAG 3432

DB 739 GTAG 763

RESULT 9

US-10-859-198-9362

; Sequence 9362, Application US/10859198

; GENERAL INFORMATION:

; APPLICANT: Wyeth

; APPLICANT: Mounts, William

; APPLICANT: Murphy, Ellen

; APPLICANT: Whitley, Maryann

; TITLE OF INVENTION: Nucleic Acid Arrays for Detecting Multiple Strains of a Non-Viral

; FILE REFERENCE: 031896-014000 (AM101085)

; CURRENT APPLICATION NUMBER: US/10/859,198

; CURRENT FILING DATE: 2004-06-03

; PRIOR APPLICATION NUMBER: US 60/475,871

; PRIOR FILING DATE: 2003-06-05

; NUMBER OF SEQ ID NOS: 282011

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 9362

; LENGTH: 4647

; TYPE: DNA

; ORGANISM: Staphylococcus aureus

US-10-859-198-9362

Query Match

1.1%; Score 48.2; DB 7; Length 4647;

Best Local Similarity 44.3%; Pred. No. 0.13;

Matches 197; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 2958 ATAAAGAGAGAGATTCGATCATGATTTAGTAGAGAGAGAGAGAGAGAGATTA 3047

DB 319 ACAAAGAGAGAGATTCGATCATGATTTAGTAGAGAGAGAGAGAGAGAGATTA 378

QY 3048 GAAAG 3107

DB 379 GAAAG 438

QY 3108 AGAAG 3167

DB 439 GAAAG 498

QY 3168 TTAAG 3227

DB 439 GAAAG 558

QY 3228 AAG 3287

DB 559 GAAAG 618

QY 3288 TTAAG 3347

DB 619 ACAAAG 678

QY 3348 CAAG 3407

DB 679 GCTGAG 738

QY 3408 GCGTTAACAAGATTTCTGCAACAG 3432

DB 739 GTAG 763

```

Oy      3168  TTAAAGCCCAAAACAAGCTGAATGATGCTGAAACAATAAAAGTAAGCAAAAACAAA 3227
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      499  GAABAGGCGCAAAAGCTGAAAGAAAACAGACAAAGCAACGAAAGAAAGCAACAAAACGAA 558
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy      3228  AAGTGGGGTCAAAAAGACAGATGTTTGTGATTCCTGCTTGATCAAAAGCTGTTGCA 3287
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      559  GAAACAAACAAAGCAACAAAGAAAGAGGCGCCAGCAAGCTGAAAGAAACAAGCAAGACGA 618
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy      3288  TTAGAAAGCGCATCTGAGGTATTATGATGCCACAGCATCGGAAAAAGATCGTACGT 3347
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      619  ACGAAGAGAGGCCCAAAAGCTGAAGAAATAGACAAAGCAACAGCAACAGAAAGGCCCAAA 678
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy      3348  CAAAGAAAGCGGAAAAACAACGCAACAAAAAATTAATTATGACCGTTATTCAAATAGT 3407
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      679  GCTAAGAAACAGAAAAAACAAGCAACAGAAAGCAACAAAAAATGAAAGAACAGACAAA 738
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy      3408  GCGTTATCAGATTATCTGCACAG 3432
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      739  GTACAAACAGAAAGAGGCCAAAG 763
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-10-859-198-9364
; Sequence 9364, Application US/10859198
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounes, William
; APPLICANT: Murphy, Ellen
; APPLICANT: Whitley, Maryann
; TITLE OF INVENTION: Nucleic Acid Arrays for Detecting Multiple Strains of a Non-Viral
; FILE REFERENCE: 031896-014000 (AM101085)

```

```

Db      679 GGTAAAGAAAACAGAAAACAGCAAGAAAGCAACCAAACTGAAGAAAACAGACAAA 738
Qy      3408 GCGTTATCAGATTATCTGCAACAG 3432
Db      739 GTAGAAACAGAGAGAGAGCGCCAAAAG 763

```

RESULT 12

```

US-10-859-198-3884
; Sequence 3884, Application US/10859198
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; APPLICANT: Murphy, Ellen
; APPLICANT: Whitely, Maryann
; TITLE OF INVENTION: Nucleic Acid Arrays for Detecting Multiple Strains of a Non-Viral
; FILE REFERENCE: 031896-014000 (AM101085)
; CURRENT APPLICATION NUMBER: US/10/859,198
; CURRENT FILING DATE: 2004-06-03
; PRIOR APPLICATION NUMBER: US 60/475,871
; PRIOR FILING DATE: 2003-06-05
; NUMBER OF SEQ ID NOS: 282011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3884
; LENGTH: 4914
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-859-198-3884

```

```

Query Match      1.0%; Score 44.6; DB 7; Length 4914;
Best Local Similarity 45.8%; Pred. No. 1.1;
Matches 192; Conservative 0; Mismatches 224; Indels 3; Gaps 1;

```

```

Qy      3017 AGTAAGCGAGAGCAAGCAAGCAATTGAAAGCCAAATGTTGAACCGACTGCTAA 3076
Db      306 AGAAAAAGCAAAATCACTGAAACAGATCAACAGAGAAAAGCATATCAACTGAAACA 365
Qy      3077 AACCAAAAGGTGA---GCCAAAAGTCGGCTCAAGAAAGACACCGAGAGCGCTTCC 3133
Db      366 AGCAACAAACAGAAAGAGCGCCAAAAGCTGAAGAAACACAAAGTGAAGAAACAGAAAGC 425
Qy      3134 TGATACCTGCTGATCAAAAGCTGTTAAAGCATTAAGACCAACAGTGAACCTGAC 3193
Db      426 GCCAAAAGCTGAAGAAACAGCAACCAAGCAACAGAGAGAGCGCCAAAAGCTGAAGAAAC 485
Qy      3194 TGCTGAAGACAAAAGTGAAGCAAAAACAAAAAAGTGGCTCAAAAAGAGCGAGTGT 3253
Db      486 AACCAAAAGCAAGAAAGAGCAACCAAAACTGAAGAAACAGCAACCAACAGAGAGA 545
Qy      3254 TTCTGATCCCTGCTGATCAAAAGCTGTTGCAATTGAAGCCGCACTTGAGGTTATGA 3313
Db      546 AGCGCCAGAGCTGAAGAAACAGCAACCAAGAGCAACAGAGAGCGCCAAAAGCTGAAGAG 605
Qy      3314 TGCCCAACAGCAATCGAAAAGATCGTCTAAGCTAAGAGAGAGCGAAAAACAGAGCA 3373
Db      606 AACCAAGCAAGCAAGCAAGAGAGAGCGCCAAAAGCTGAAGAAACAGAAAAACAGCAAC 665
Qy      3374 AAAAAAGACTGATGAGCGGTTATTCATATAGTCGTTATCAATATATCTGCAAGAG 3432
Db      666 AGAAGAGCAACCAAAAGCTGAAGAAAGCAAGAGTGAAGAAACAGAGAGCGCCAAAAG 724

```

RESULT 13

```

US-10-859-198-11736
; Sequence 11736, Application US/10859198
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; APPLICANT: Murphy, Ellen
; APPLICANT: Whitely, Maryann
; TITLE OF INVENTION: Nucleic Acid Arrays for Detecting Multiple Strains of a Non-Viral

```

```

; TITLE OF INVENTION: Species
; FILE REFERENCE: 031896-014000 (AM101085)
; CURRENT APPLICATION NUMBER: US/10/859,198
; CURRENT FILING DATE: 2004-06-03
; PRIOR APPLICATION NUMBER: US 60/475,871
; PRIOR FILING DATE: 2003-06-05
; NUMBER OF SEQ ID NOS: 282011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11736
; LENGTH: 4914
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-859-198-11736

```

```

Query Match      1.0%; Score 44.6; DB 7; Length 4914;
Best Local Similarity 45.8%; Pred. No. 1.1;
Matches 192; Conservative 0; Mismatches 224; Indels 3; Gaps 1;

```

```

Qy      3017 AGTAAGCGAGAGCAAGCAAGCAATTGAAAGCCAAATGTTGAACCGACTGCTAA 3076
Db      306 AGAAAAAGCAAAATCACTGAAACAGATCAACAGAGAAAAGCATATCAACTGAAACA 365
Qy      3077 AACCAAAAGGTGA---GCCAAAAGTCGGCTCAAGAAAGACACCGAGAGCGCTTCC 3133
Db      366 AGCAACAAACAGAAAGAGCGCCAAAAGCTGAAGAAACACAAAGTGAAGAAACAGAAAGC 425
Qy      3134 TGATACCTGCTGATCAAAAGCTGTTAAAGCATTAAGACCAACAGTGAACCTGAC 3193
Db      426 GCCAAAAGCTGAAGAAACAGCAACCAAGAGCAACAGAGAGCGCCAAAAGCTGAAGAAAC 485
Qy      3194 TGCTGAAGACAAAAGTGAAGCAAAAACAAAAAAGTGGCTCAAAAAGAGCGAGTGT 3253
Db      486 AACCAAAAGCAAGAAAGAGCAACCAAAACTGAAGAAACAGCAACCAACAGAGAGA 545
Qy      3254 TTCTGATCCCTGCTGATCAAAAGCTGTTGCAATTGAAGCCGCACTTGAGGTTATGA 3313
Db      546 AGCGCCAGAGCTGAAGAAACAGCAACCAAGAGCAACAGAGAGCGCCAAAAGCTGAAGAG 605
Qy      3314 TGCCCAACAGCAATCGAAAAGATCGTCTAAGCTAAGAGAGAGCGAAAAACAGAGCA 3373
Db      606 AACCAAGCAAGCAAGCAAGAGAGAGCGCCAAAAGCTGAAGAAACAGAAAAACAGCAAC 665
Qy      3374 AAAAAAGACTGATGAGCGGTTATTCATATAGTCGTTATCAATATATCTGCAAGAG 3432
Db      666 AGAAGAGCAACCAAAAGCTGAAGAAAGCAAGAGTGAAGAAACAGAGAGCGCCAAAAG 724

```

RESULT 14

```

PCT-US03-41761-38500
; Sequence 38500, Application PCT/US0341761
; GENERAL INFORMATION:
; APPLICANT: MWI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: CHARTERIS, Paul
; APPLICANT: ROSENFIELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; TITLE OF INVENTION: COMPOSITIONS, METHODS, AND SYSTEMS FOR INFERRING BOVINE BREED
; FILE REFERENCE: MW111500
; CURRENT APPLICATION NUMBER: PCT/US03/41761
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38500
; LENGTH: 1543
; TYPE: DNA
; ORGANISM: Bovine
PCT-US03-41761-38500

```

```

Query Match      1.0%; Score 44.4; DB 1; Length 1543;
Best Local Similarity 50.0%; Pred. No. 0.77;

```


Matches 111; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

```
QY 3800 TCATCGAAGGCGATTAATTAAGCGGAAATGCAAGTATCATGCTTCGTTAGGGCAATT 3859
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1068 TCATCGAAGGCGATTAATTAAGGTTGAGTGAAGTATTAAGGCTGAGATTTT 1127
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3860 GGGCAATTCAGCTTATTTTGAGTTAATGCTATTTTATTAAGTGAATTAATCAATC 3919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1128 CTGATTAATCTGCTGATTTCTGATTAATTTCTGATTTTCTGATTAATTAATTTT 1187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3920 TGAGGAGTGAAGTGAAGGCGCTGACCTTGATTTAATGCTATTAATGCTGCAATCG 3979
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1188 ATAAAGATTAAAGAAAGCTCTATAGAACTATAGTAAGTACGACGAGAGAAATTA 1247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3980 AGTGATTATACATTTACTCGACGATTAATATCAGCGTTAA 4021
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1248 ACTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 15

PCT-US03-41761-38590
; Sequence 38500, Application PC/TUS0341761
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISER, Sue K.
; APPLICANT: CHARTERIS, Paul
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; TITLE OF INVENTION: COMPOSITIONS, METHODS, AND SYSTEMS FOR INFERRING BOVINE BREED
; FILE REFERENCE: NM1150M0
; CURRENT APPLICATION NUMBER: PCT/US03/41761
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent version 3.1
; SEQ ID NO 38500
; LENGTH: 1543
; TYPE: DNA
; ORGANISM: Bovine 19866801025522
PCT-US03-41761-38500

Query Match 1.0%; Score 44.4; DB 1; Length 1543;

Best Local Similarity 50.0%; Pred. No. 0.77;

Matches 111; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

```
QY 3800 TCATCGAAGGCGATTAATTAAGCGGAAATGCAAGTATCATGCTTCGTTAGGGCAATT 3859
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1068 TCATCGAAGGCGATTAATTAAGGTTGAGTGAAGTATTAAGGCTGAGATTTT 1127
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3860 GGGCAATTCAGCTTATTTTGAGTTAATGCTATTTTATTAAGTGAATTAATCAATC 3919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1128 CTGATTAATCTGCTGATTTCTGATTAATTTCTGATTTTCTGATTAATTAATTTT 1187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3920 TGAGGAGTGAAGTGAAGGCGCTGACCTTGATTTAATGCTATTAATGCTGCAATCG 3979
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1188 ATAAAGATTAAAGAAAGCTCTATAGAACTATAGTAAGTACGACGAGAGAAATTA 1247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3980 AGTGATTATACATTTACTCGACGATTAATATCAGCGTTAA 4021
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1248 ACTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: June 24, 2004, 19:55:42
Job time : 670 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2004, 08:14:26 ; Search time 275 Seconds
(without alignments)
8715.756 Million cell updates/sec

Title: US-10-645-655-1
Perfect score: 4319
Sequence: 1 TCATAGTCGTTACTACTGTTTATTTATTCGAAATTAATTAACCTTA 4319

Scoring table: IDENTITY NUC
Gapop 10.0, Gapept 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4319	100.0	4319	3	US-08-296-791-1
2	4319	100.0	4319	4	US-09-839-996-1
3	4319	100.0	4319	5	US-10-080-505-1
4	4319	100.0	4319	5	PCT-US95-10661A-1
5	3180.8	73.6	4305	4	US-10-080-505-8
6	3019.8	69.9	1830121	4	US-09-557-884-1
7	3019.8	69.9	1830121	4	US-09-643-990A-1
8	2990.4	69.2	5245	4	US-10-080-505-12
9	2970	68.8	4828	4	US-10-080-505-16
10	2770.2	64.1	4822	4	US-10-080-505-14
11	2709.6	62.7	4605	4	US-10-080-505-10
12	164.8	3.8	4899	2	US-08-210-535-5
13	164.8	3.8	4899	6	5268270-1
14	62.2	1.4	7218	1	US-08-232-463-14
15	48.6	1.1	832	4	US-08-621-976-2813
16	44.6	1.0	1664976	4	US-08-916-421B-1
17	44.2	1.0	3312	4	US-09-601-198-58
18	42.8	1.0	5610	4	US-10-204-708-54
19	42.8	1.0	11091	4	US-09-134-001C-2243
20	42.6	1.0	3343	3	US-08-965-762-19
21	42.6	1.0	3343	4	US-09-911-927-19
22	42.6	1.0	3343	4	US-09-911-927-21
23	42.6	1.0	3343	4	US-09-911-882-19
24	42.6	1.0	3343	4	US-09-911-882-21
25	42.6	1.0	3343	4	US-09-911-888-19
26	42.6	1.0	3343	4	US-09-911-888-21
27	42.6	1.0	580073	4	US-08-545-528D-1

28	41.6	1.0	832	4	US-09-621-976-2813	Sequence 2813, App1
29	40.4	0.9	399	4	US-09-621-976-8976	Sequence 8976, App1
30	40.2	0.9	1664976	4	US-08-916-421B-1	Sequence 1, App1
31	39.4	0.9	2226	4	US-09-543-681A-1364	Sequence 1364, App1
32	39.4	0.9	4770	4	US-09-543-681A-826	Sequence 826, App1
33	39.4	0.9	1830121	4	US-09-557-884-1	Sequence 1, App1
34	39.4	0.9	1830121	4	US-09-643-990A-1	Sequence 222, App1
35	39.2	0.9	3001	4	US-09-539-333D-223	Sequence 223, App1
36	39.2	0.9	3001	4	US-09-539-333D-224	Sequence 224, App1
37	39.2	0.9	3001	4	US-10-204-708-52	Sequence 52, App1
38	39	0.9	5219	4	US-09-601-198-60	Sequence 1516, App1
39	38.4	0.9	1197	4	US-09-134-000C-1518	Sequence 1518, App1
40	37.8	0.9	387	4	US-09-621-976-15562	Sequence 15562, App1
41	37.8	0.9	771	4	US-09-543-681A-4041	Sequence 4041, App1
42	37.8	0.9	640681	4	US-09-790-988-1	Sequence 1, App1
43	37.6	0.9	1037	4	US-09-181-585-3	Sequence 1, App1
44	37.6	0.9	1159	4	US-09-181-585-1	Sequence 1, App1
45	37.6	0.9	1159	4	US-09-181-585-1	Sequence 1, App1

ALIGNMENTS

RESULT 1
US-08-296-791-1
Sequence 1, Application US/08296791
Patent No. 6245337
GENERAL INFORMATION:
APPLICANT: St. Geme III, Joseph W.
APPLICANT: Falkow, Stanley
TITLE OF INVENTION: Haemophilus Adherence and Penetration
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Flehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/296,791
FILING DATE: 25-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Treacart, Richard F.
REGISTRATION NUMBER: 31,801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1983
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4319 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
FEATURE:
NAME/KEY: CDS
LOCATION: 60..4241
US-08-296-791-1
Query Match: 100.0%; Score 4319; DB 3; Length 4319;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCATAGTCGTTACTACTGTTTATTTATTCGAAATTAATTAACCTTA 60

Db 1 TCATAGTCGTTTACAGATTTTATATACAAAATTTCTAATTAATAAATCAATTA 60
 Qy 61 TGAATAAAACGTAATTCGCTTAATTTTAAACCGCTGATTTCTATAGGAATGAT 120
 Db 61 TGAATAAAACGTAATTCGCTTAATTTTAAACCGCTGATTTCTATAGGAATGAT 120
 Qy 121 CGCAAGCGTGGCTGGTCACTTATTTTGGATTTGATTAACATATATATGATTTG 180
 Db 121 CGCAAGCGTGGCTGGTCACTTATTTTGGATTTGATTTACCATATATATGATTTG 180
 Qy 181 CCGAGATTAAGGGAAGTTCAAGTGGGGCTCAAAATATTAGTTTAAACAAACAG 240
 Db 181 CCGAGATTAAGGGAAGTTCAAGTGGGGCTCAAAATATTAGTTTAAACAAACAG 240
 Qy 241 GCGAATTTAGTGGCAATCAATGACAAACCCCAATGATTTTCTGTAAGTCA 300
 Db 241 GCGAATTTAGTGGCAATCAATGACAAACCCCAATGATTTTCTGTAAGTCA 300
 Qy 301 GTAAAGCGTGGAGCGCTTGTAATAATCATATATGATGCGTGGCAATTAAGT 360
 Db 301 GTAAAGCGTGGAGCGCTTGTAATAATCATATATGATGCGTGGCAATTAAGT 360
 Qy 361 GATATACGATTTGATTTTGGTGAAGGGAACACCCGATCAATCGTTTACTT 420
 Db 361 GATATACGATTTGATTTTGGTGAAGGGAACACCCGATCAATCGTTTACTT 420
 Qy 421 ATAAGATTGTAAGAAATCAATCAAAATTAATTTATCATCTTATGAGGACAT 480
 Db 421 ATAAGATTGTAAGAAATCAATCAAAATTAATTTATCATCTTATGAGGACAT 480
 Qy 481 ACCATTAATCAAGATTAATCAATTCGTTACAGAAAGCGCTCGAATGATATG 540
 Db 481 ACCATTAATCAAGATTAATCAATTCGTTACAGAAAGCGCTCGAATGATATG 540
 Qy 541 ATATGATGCGAGTACTTATTCAGATGAACAAATATCAAGACGTTTCTATCG 600
 Db 541 ATATGATGCGAGTACTTATTCAGATGAACAAATATCAAGACGTTTCTATCG 600
 Qy 601 CTGACGAGCGATTTTGGGAAATGATCAACAAAGGCAACAGTTGCGGTGATAC 660
 Db 601 CTGACGAGCGATTTTGGGAAATGATCAACAAAGGCAACAGTTGCGGTGATAC 660
 Qy 661 ATATCTGACAGCTGGCAATACACATCAAGCTGAGAGCAATGATATGATTT 720
 Db 661 ATATCTGACAGCTGGCAATACACATCAAGCTGAGAGCAATGATATGATTT 720
 Qy 721 TGGAGGCGATTTGTTGAAGCGGAGATATGCTCATTCGATTCGAGCTCAAG 780
 Db 721 TGGAGGCGATTTGTTGAAGCGGAGATATGCTCATTCGATTCGAGCTCAAG 780
 Qy 781 GGGACAGTGTCTCGAATGTTATTTATGATGCTGAAGAAACAAATATGTTAT 840
 Db 781 GGGACAGTGTCTCGAATGTTATTTATGATGCTGAAGAAACAAATATGTTAT 840
 Qy 841 GGAATATTACGGGAAGCAACCTTTGAAGGCAAGAAATGGGTTCAATGTTGCA 900
 Db 841 GGAATATTACGGGAAGCAACCTTTGAAGGCAAGAAATGGGTTCAATGTTGCA 900
 Qy 901 AATCTTATTTTGAAGAAATTTTGAAGGATTTATCATATCACTTCAACCCAG 960
 Db 901 AATCTTATTTTGAAGAAATTTTGAAGGATTTATCATATCACTTCAACCCAG 960
 Qy 961 GTAAAGAGTGTACAAATTAAGTGAATGATGATGAGGGCTATATCACTGAAT 1020
 Db 961 GTAAAGAGTGTACAAATTAAGTGAATGATGATGAGGGCTATATCACTGAAT 1020
 Qy 1021 CAGGAATCCATGAGAAATTAATTAATGATGAGCAATATGATTTACCTTTGAAG 1080
 Db 1021 CAGGAATCCATGAGAAATTAATTAATGATGAGCAATATGATTTACCTTTGAAG 1080
 Qy 1081 AGGATTAAGTCTAATCTGATATGAGAGCTAATATTTATCTCGAGTTTAAAC 1140
 Db 1081 AGGATTAAGTCTAATCTGATATGAGAGCTAATATTTATCTCGAGTTTAAAC 1140
 Qy 1281 AGGATTAAGTCTAATCTGATATGAGAGCTAATATTTATCTCGAGTTTAAAC 1140
 Db 1281 AGGATTAAGTCTAATCTGATATGAGAGCTAATATTTATCTCGAGTTTAAAC 1140
 Qy 1441 ATGAGAAAGCGCTAATTTATGATGACAAACAAAGATATTAATCTCGACATG 1200
 Db 1441 ATGAGAAAGCGCTAATTTATGATGACAAACAAAGATATTAATCTCGACATG 1200
 Qy 1201 TTAACCAAGGCGGGTGTCTTATTTTGAAGGATTTTACAGATCTCAATTTCA 1260
 Db 1201 TTAACCAAGGCGGGTGTCTTATTTTGAAGGATTTTACAGATCTCAATTTCA 1260
 Qy 1261 ACCAACTTGGCAAGAGCTGGCATATGATGATGAAATGACACCGTTACTTGA 1320
 Db 1261 ACCAACTTGGCAAGAGCTGGCATATGATGATGAAATGACACCGTTACTTGA 1320
 Qy 1321 TAAATGCGTGGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 Db 1321 TAAATGCGTGGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 Qy 1381 CCAAGGCGGAAATTAAGTTCGATCAGGAGGAGGATGATGATGATGATGATGAT 1440
 Db 1381 CCAAGGCGGAAATTAAGTTCGATCAGGAGGAGGATGATGATGATGATGATGAT 1440
 Qy 1441 AGGACAGATCAAGGCAACAAAGCGCTTATGATGATGATGATGATGATGAT 1500
 Db 1441 AGGACAGATCAAGGCAACAAAGCGCTTATGATGATGATGATGATGATGAT 1500
 Qy 1501 GGAATGCGTGAATTAAGGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 Db 1501 GGAATGCGTGAATTAAGGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 Qy 1561 GTGATGCTGCTGATGATTTAAGGCAATCAATTAATTTAAGGATTTAAGGAT 1620
 Db 1561 GTGATGCTGCTGATGATTTAAGGCAATCAATTAATTTAAGGATTTAAGGAT 1620
 Qy 1621 ACAGAGGCGCAATGATTTGATGATGATGATGATGATGATGATGATGATGAT 1680
 Db 1621 ACAGAGGCGCAATGATTTGATGATGATGATGATGATGATGATGATGATGAT 1680
 Qy 1681 GGAAGGAAAGGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 Db 1681 GGAAGGAAAGGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 Qy 1741 AATATGCTCAAGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 1800
 Db 1741 AATATGCTCAAGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 1800
 Qy 1801 ACTTATTTTAAACCAACCAAGGATGATGATGATGATGATGATGATGATGAT 1860
 Db 1801 ACTTATTTTAAACCAACCAAGGATGATGATGATGATGATGATGATGATGAT 1860
 Qy 1861 TAAAGGCGATTTACCAACCAAGGATGATGATGATGATGATGATGATGATGAT 1920
 Db 1861 TAAAGGCGATTTACCAACCAAGGATGATGATGATGATGATGATGATGATGAT 1920
 Qy 1921 AGGCTCAACATCAATTTAATTAAGCTGATGATGATGATGATGATGATGATGAT 1980
 Db 1921 AGGCTCAACATCAATTTAATTAAGCTGATGATGATGATGATGATGATGATGAT 1980
 Qy 1981 TTTGTTGGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 Db 1981 TTTGTTGGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 Qy 2041 GCGAAGTGGGCTGTTTCTCGAATGTTTCTCAATGAGGAAATGAGCACTGAGCA 2100
 Db 2041 GCGAAGTGGGCTGTTTCTCGAATGTTTCTCAATGAGGAAATGAGCACTGAGCA 2100
 Qy 2101 ATATGCAATGCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2160
 Db 2101 ATATGCAATGCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2160
 Qy 2161 CAGATTTGAGGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2220
 Db 2161 CAGATTTGAGGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2220

QY	2221	ATTCTAATCAAAAACAAATCAATGCTCATATATTTATTTATCTGATATGCAACGGCA	2283
Db	2221	ATTCTAATCAAAAACAAATCAATGCTCATATATTTATTTATCTGATATGCAACGGCA	2283
QY	2281	ATGTTAAAGTTTACAAAATTAAATGCAATGTCACTTTAACAAATCAACGCCATTTA	2340
Db	2281	ATGTTAAAGTTTACAAAATTAAATGCAATGTCACTTTAACAAATCAACGCCATTTA	2340
QY	2341	CATTAAAGCAAAATGCCAATGAGCAATATTCGACTTCGCGCAATTCATCGTCA	2400
Db	2341	CATTAAAGCAAAATGCCAATGAGCAATATTCGACTTCGCGCAATTCATCGTCA	2400
QY	2401	CGGTGATATATCAAACTTTGCAACGGTATGTGATTTAAACGATCAATTTCTT	2460
Db	2401	CGGTGATATATCAAACTTTGCAACGGTATGTGATTTAAACGATCAATTTCTT	2460
QY	2461	TAAAAAACGCCATTTTTCGCAACMAATTCGAGGACAAAGGACAAACAGTGCCTGG	2520
Db	2461	TAAAAAACGCCATTTTTCGCAACMAATTCGAGGACAAAGGACAAACAGTGCCTGG	2520
QY	2521	AAATATGCACTGCGCAATTCCTAGCAATCTACATTCGCAATTTAAGCTAAATACA	2580
Db	2521	AAATATGCACTGCGCAATTCCTAGCAATCTACATTCGCAATTTAAGCTAAATACA	2580
QY	2581	GTACGATCAAGTTAAATTCAGCTTATTCAGCTACCTCAACATATAGCGACGTGCGGTT	2640
Db	2581	GTACGATCAAGTTAAATTCAGCTTATTCAGCTACCTCAACATATAGCGACGTGCGGTT	2640
QY	2641	CATTAGAGCGGAAACCAAGCCAAATGCGCAACAATGTTTCAACATTTGACGTAA	2700
Db	2641	CATTAGAGCGGAAACCAAGCCAAATGCGCAACAATGTTTCAACATTTGACGTAA	2700
QY	2701	ATGTTAAATTTAGTGGGCAACACATATTCATTTACTTCATTTATTTGGCATTAATA	2760
Db	2701	ATGTTAAATTTAGTGGGCAACACATATTCATTTACTTCATTTATTTGGCATTAATA	2760
QY	2761	GGCATTAATTAATAATTTCATATGACGTGAGGGCGATTCATATTTATCTGTTGCAACA	2820
Db	2761	GGCATTAATTAATAATTTCATATGACGTGAGGGCGATTCATATTTATCTGTTGCAACA	2820
QY	2821	CAGCAAAACAAACCGAAACCCCTTGAGCAATTAATCTTGGTGAACCAATATATCAAC	2880
Db	2821	CAGCAAAACAAACCGAAACCCCTTGAGCAATTAATCTTGGTGAACCAATATATCAAC	2880
QY	2881	CGTTATCAGATAGCTCAAAATTTACTTTAGAAAATGACACGCTTGAAGCAGTGCATTAC	2940
Db	2881	CGTTATCAGATAGCTCAAAATTTACTTTAGAAAATGACACGCTTGAAGCAGTGCATTAC	2940
QY	2941	GTATTAATTTAGTGAATATGATGAGCGAATTCGCTGTCATTAACCCATTAATAAGAGAG	3000
Db	2941	GTATTAATTTAGTGAATATGATGAGCGAATTCGCTGTCATTAACCCATTAATAAGAGAG	3000
QY	3001	AATTCACATATGATTTTGTATAGAGGACGAAACGAAACGAACTTTAGACGAAACAAG	3060
Db	3001	AATTCACATATGATTTTGTATAGAGGACGAAACGAAACGAACTTTAGACGAAACAAG	3060
QY	3061	TTGAAACGACCTCTTAAACCAAAACAGGTGAGCAAAAGTCCGCTCAAGAAAGACAGCA	3120
Db	3061	TTGAAACGACCTCTTAAACCAAAACAGGTGAGCAAAAGTCCGCTCAAGAAAGACAGCA	3120
QY	3121	GAGAGAGGTTTCTGATATACCTGCTGATCAAAAGCTGTTTAAACGCAATTAAGACCAAC	3180
Db	3121	GAGAGAGGTTTCTGATATACCTGCTGATCAAAAGCTGTTTAAACGCAATTAAGACCAAC	3180
QY	3181	AAGCTGAATGACTGCTGAACCAAAAAGTAAAGGCAAAAACAAAAGTGGCTCA	3240
Db	3181	AAGCTGAATGACTGCTGAACCAAAAAGTAAAGGCAAAAACAAAAGTGGCTCA	3240
QY	3241	AAAGACAGCTGTTTCTGATATCCCTGCTGATCAAAAGCTGTGATTAAGACCGAC	3300
Db	3241	AAAGACAGCTGTTTCTGATATCCCTGCTGATCAAAAGCTGTGATTAAGACCGAC	3300

QY	3301	TTGAGGTTTGGATGCCCAACGAAATCGGAAAAGAGTCGTAGCTCAAAAGACGG	3360
Db	3301	TTGAGGTTTGGATGCCCAACGCAATCGGAAAAGAGTCGTAGCTCAAAAGACGG	3360
QY	3361	AAAAACAACGCAACAAAAAGCTTGAACAGCGCTTATTCAAATAGTCGTTACAGAT	3420
Db	3361	AAAAACAACGCAACAAAAAGCTTGAACAGCGCTTATTCAAATAGTCGTTACAGAT	3420
QY	3361	AAAAACAACGCAACAAAAAGCTTGAACAGCGCTTATTCAAATAGTCGTTACAGAT	3420
Db	3361	AAAAACAACGCAACAAAAAGCTTGAACAGCGCTTATTCAAATAGTCGTTACAGAT	3420
QY	3421	TATCGCAACGTAATAAGATAGCTTCTCTGCAGATGAATTAAGATCGTCTTTTGTAG	3480
Db	3421	TATCGCAACGTAATAAGATAGCTTCTCTGCAGATGAATTAAGATCGTCTTTTGTAG	3480
QY	3481	ATCAACACAAATCTGCCGTGTGACAAATATGCGACAGAGTAATAACAGCTATCTTCG	3540
Db	3481	ATCAACACAAATCTGCCGTGTGACAAATATGCGACAGAGTAATAACAGCTATCTTCG	3540
QY	3541	ATGGTTCGGCTGCTTAACAGCAAGCAACCACTTAACGTTACAACTGGGGTGCAAAAG	3600
Db	3541	ATGGTTCGGCTGCTTAACAGCAAGCAACCACTTAACGTTACAACTGGGGTGCAAAAG	3600
QY	3601	CTTTAGCTTAATGACGAATTTGGACAGCTTTCTTCGATAGCGCTTCAGATATCTCTTG	3660
Db	3601	CTTTAGCTTAATGACGAATTTGGACAGCTTTCTTCGATAGCGCTTCAGATATCTCTTG	3660
QY	3661	ATGACAGGTTTAAAAATCAAGCAATCAACATGATGCGGCTTTGGCCATATCAAT	3720
Db	3661	ATGACAGGTTTAAAAATCAAGCAATCAACATGATGCGGCTTTGGCCATATCAAT	3720
QY	3721	GGGGCGATTACAATTGGTGTAAACGTGGGAACGGAACTAGTCGAGTAATAATGACTG	3780
Db	3721	GGGGCGATTACAATTGGTGTAAACGTGGGAACGGAACTAGTCGAGTAATAATGACTG	3780
QY	3781	AAGACAAAGCGCAAAATTTCAATCGAAAAACGATAAATTAGCGGAATCAAGTATTC	3840
Db	3781	AAGACAAAGCGCAAAATTTCAATCGAAAAACGATAAATTAGCGGAATCAAGTATTC	3840
QY	3841	AGTTCCGTTTAGGGCAATTTGGGCACTCAGCGTTATTTGGAGTTAATGCTATTTATTTG	3900
Db	3841	AGTTCCGTTTAGGGCAATTTGGGCACTCAGCGTTATTTGGAGTTAATGCTATTTATTTG	3900
QY	3961	AAGGTAAATTTCAATCAATCTGAGAACTGAGATGAGTAAAGAGGCTACCTGCAATTTATC	3960
Db	3961	AAGGTAAATTTCAATCAATCTGAGAACTGAGATGAGTAAAGAGGCTACCTGCAATTTATC	3960
QY	4021	AGGCTATTTCTTGTCATTTAAGTATAGTTTGAACGGTAAACGTAACCAACGGTAA	4080
Db	4021	AGGCTATTTCTTGTCATTTAAGTATAGTTTGAACGGTAAACGTAACCAACGGTAA	4080
QY	4081	ATTCACAGGTGTGCAACACCAATTTGACGTTATTTGCAAAAAGAGTGGGATTAAAG	4140
Db	4081	ATTCACAGGTGTGCAACACCAATTTGACGTTATTTGCAAAAAGAGTGGGATTAAAG	4140
QY	4141	CAGAAATTTTCAATTCGAAATTTCCGCTTTATTCGAAATTCGAAAGTTCGAACTCG	4200
Db	4141	CAGAAATTTTCAATTCGAAATTTCCGCTTTATTCGAAATTCGAAAGTTCGAACTCG	4200
QY	4201	GCAACAGCAAAATGTGGCGTGAATTTGGGCTATCGTGGTAAATAACACATAATTT	4260
Db	4201	GCAACAGCAAAATGTGGCGTGAATTTGGGCTATCGTGGTAAATAACACATAATTT	4260
QY	4261	ATGCTTTATATGAATAACAGAGGTGGGTCAATCAATCCACCTTTTATTCGAAATAT	4319
Db	4261	ATGCTTTATATGAATAACAGAGGTGGGTCAATCAATCCACCTTTTATTCGAAATAT	4319

RESULT 2
US-09-839-996-1
: Sequence 1, Application US/09839996
: Patent No. 6642371

Db 1441 AGGAGAGATCAAGGCAACAAAGCTTTAGTAAATGGCTTGGTAAAGGCAAG 1500
Qy 1501 GGAATGTTCAATTAACATGATTAACATTTGATACGATTAATTTATTTATTTGCTTC 1560
Db 1501 GGAATGTTCAATTAACATGATTAACATTTGATACGATTAATTTATTTATTTGCTTC 1560
Qy 1561 GTGGTGGTCTTGAATCTTAAAGGCAATCATTAACCTTTAAAGTATCCAAATACGG 1620
Db 1561 GTGGTGGTCTTGAATCTTAAAGGCAATCATTAACCTTTAAAGTATCCAAATACGG 1620
Qy 1621 AGAGAGGAGCAATGATTTGGAACCAATCACTGAAGCCGCTAATGATCATTTATCTG 1680
Db 1621 AGAGAGGAGCAATGATTTGGAACCAATCACTGAAGCCGCTAATGATCATTTATCTG 1680
Qy 1681 GGAAGCAAGATTTGTTACCTAATGGAATTAATTAATTAATGATTAACGAAG 1740
Db 1681 GGAAGCAAGATTTGTTACCTAATGGAATTAATTAATTAATGATTAACGAAG 1740
Qy 1741 AAATGGCTTCAACAGGTTGGTTGGCAAAACAGATTAATAATTAACATGGGCAATTA 1800
Db 1741 AAATGGCTTCAACAGGTTGGTTGGCAAAACAGATTAATAATTAACATGGGCAATTA 1800
Qy 1801 ACCATTTTAAACCAACCAAGAGATGATCTTGCTAAGGTTGATGATTAATTT 1860
Db 1801 ACCATTTTAAACCAACCAAGAGATGATCTTGCTAAGGTTGATGATTAATTT 1860
Qy 1861 TAAAGGCGATTTATCCCAACAAAGGTAACTATTTTCAAGGCGTAAAGCAAGCCG 1920
Db 1861 TAAAGGCGATTTATCCCAACAAAGGTAACTATTTTCAAGGCGTAAAGCAAGCCG 1920
Qy 1921 AGGCTTCAATCATTTAAATTAACGTTGGTCAAGAAATGGAAGGTATCAACAGGCAAA 1980
Db 1921 AGGCTTCAATCATTTAAATTAACGTTGGTCAAGAAATGGAAGGTATCAACAGGCAAA 1980
Qy 1981 TTGGTGGGATCAAGTTGATCAACGATTAATTAAGGTAAGGTAAGGTAAGGTAAG 2040
Db 1981 TTGGTGGGATCAAGTTGATCAACGATTAATTAAGGTAAGGTAAGGTAAGGTAAG 2040
Qy 2041 GCGAGAGTGGGTTGTTCTCGCAATGTTCTTCAATTAAGGTAAGGTAAGGTAAG 2100
Db 2041 GCGAGAGTGGGTTGTTCTCGCAATGTTCTTCAATTAAGGTAAGGTAAGGTAAG 2100
Qy 2101 AATATGCAATGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 2160
Db 2101 AATATGCAATGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 2160
Qy 2161 CAGATGGAAGATTAACGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2220
Db 2161 CAGATGGAAGATTAACGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2220
Qy 2221 ATTCTATACCAAAACAAATCAATGATTTGATTTGATTTGATTTGATTTGATTTGAT 2280
Db 2221 ATTCTATACCAAAACAAATCAATGATTTGATTTGATTTGATTTGATTTGATTTGAT 2280
Qy 2281 ATGTTAAAGTTTGAAGAACTTAATGCAATGCACTTTAAACAAATCAAGCAATTTA 2340
Db 2281 ATGTTAAAGTTTGAAGAACTTAATGCAATGCACTTTAAACAAATCAAGCAATTTA 2340
Qy 2341 CATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2400
Db 2341 CATTAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2400
Qy 2401 CGGTGATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2460
Db 2401 CGGTGATTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2460
Qy 2461 TAAATAACAGCAATTTTGAAGCAATTTGATTTGATTTGATTTGATTTGATTTGAT 2520
Db 2461 TAAATAACAGCAATTTTGAAGCAATTTGATTTGATTTGATTTGATTTGATTTGAT 2520
Qy 2521 AAAATGCACTTGAAGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2580
Db 2521 AAAATGCACTTGAAGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2580

Db 2521 AAAATGCACTTGAAGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2580
Qy 2581 GTACGATCAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2640
Db 2581 GTACGATCAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2640
Qy 2641 CATTAAGCAAGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2700
Db 2641 CATTAAGCAAGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2700
Qy 2701 ATGTTAAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2760
Db 2701 ATGTTAAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2760
Qy 2761 GCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2820
Db 2761 GCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2820
Qy 2821 CAGGCAAGCAAGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2880
Db 2821 CAGGCAAGCAAGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2880
Qy 2881 CGTTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2940
Db 2881 CGTTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2940
Qy 2941 GTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3000
Db 2941 GTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3000
Qy 3001 AATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3060
Db 3001 AATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3060
Qy 3061 TTGAACGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3120
Db 3061 TTGAACGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3120
Qy 3121 GAGGAGGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 3180
Db 3121 GAGGAGGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 3180
Qy 3181 AAGCTGAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3240
Db 3181 AAGCTGAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3240
Qy 3241 AAGGAGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3300
Db 3241 AAGGAGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3300
Qy 3301 TTGAGGTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3360
Db 3301 TTGAGGTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3360
Qy 3361 AAAAAGCAAGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3420
Db 3361 AAAAAGCAAGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3420
Qy 3421 TATCTGAACAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3480
Db 3421 TATCTGAACAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3480
Qy 3481 ATCAAGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3540
Db 3481 ATCAAGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3540
Qy 3541 ATGCGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3600
Db 3541 ATGCGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3600
Qy 3601 CTTAGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 3660
Db 3601 CTTAGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 3660

Qy	3661	ATGAACAGGTTAAATAATACGACGATTAACGATGATGCGGGTTTGCCAAATTCAT	3720
Ds	3661	ATGACAGGTTAAATAATACGACGACATTAACGATGATGCGGGTTTGCCAAATTCAT	3720
Qy	3721	GGGGCGATTTACATTTGGTGTAAACGTGGGAACGGAAATAGTCCGATTAATGACTG	3780
Ds	3721	GGGGCGATTTCAATTTGGTGTAAACGTGGGAACGGAAATAGTCCGATTAATGACTG	3780
Qy	3781	AAGAATAAGCGCAAAAATTTCAATCGAAAACCGTAAATATATGCGCGAATCGAATATC	3840
Ds	3781	AAGAATAAGCGCAAAAATTTCAATCGAAAACCGTAAATATATGCGCGAATCGAATATC	3840
Qy	3841	AGTTCCGTTTAGCGCAATTGGGCAATCGACCTATTTTGGAGTTATGCGATATTTATG	3900
Ds	3841	AGTTCCGTTTAGCGCAATTGGGCAATCGACCTATTTTGGAGTTATGCGATATTTATG	3900
Qy	3901	AACGTAAAAATTAATCAATCTGAGAGAGTAGAGTAGAATAACCGCTACCTTCATTAATC	3960
Ds	3901	AACGTAAAAATTAATCAATCTGAGAGAGTAGAGTAGAATAACCGCTACCTTCATTAATC	3960
Qy	3961	GCATATAAGTCGGGATTCGAGTATATATACATTTATCTCCGACAGATATATACAGGTTA	4020
Ds	3961	GCATATAAGTCGGGATTCGAGTATATATACATTTATCTCCGACAGATATATACAGGTTA	4020
Qy	4021	AGCCTTAATTCCTGCTCAATATATGTTATGTTCAAAACGTAAACGTAAACCAACGTTAA	4080
Ds	4021	AGCCTTAATTCCTGCTCAATATATGTTATGTTCAAAACGTAAACGTAAACCAACGTTAA	4080
Qy	4081	ATCTCACGCTGTTCGACAAACCAATTTGGACGTTATTTGGCAAAAAGAGTGGATTAAG	4140
Ds	4081	ATCTCACGCTGTTCGACAAACCAATTTGGACGTTATTTGGCAAAAAGAGTGGATTAAG	4140
Qy	4141	CAGAAATTTATCAATTCGAAATTTCCGCTTTATCTCAAAATCTCAAGGTCACAACTCG	4200
Ds	4141	CAGAAATTTATCAATTCGAAATTTCCGCTTTATCTCAAAATCTCAAGGTCACAACTCG	4200
Qy	4201	GCAAAACGCAAAATGTGGCGTGAATTTGCGCTATCTGTGGTAAATAATCAACATTAATTT	4260
Ds	4201	GCAAAACGCAAAATGTGGCGTGAATTTGCGCTATCTGTGGTAAATAATCAACATTAATTT	4260
Qy	4261	ATGCTTATATGATTAACAGAGTGGGTGCAGATCGAATCCACTTTTATTCGAATAT	4320
Ds	4261	ATGCTTATATGATTAACAGAGTGGGTGCAGATCGAATCCACTTTTATTCGAATAT	4320

```

RESULT 3--
TS-10-030-505-1
? Sequence 1, Application US/10080505
? Patent No. 6576948
? GENERAL INFORMATION:
? APPLICANT: Sc. Geme, Joseph W.
? TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINMS
? FILE REFERENCE: A-59941-1/REF/DCE/IDR
? CURRENT APPLICATION NUMBER: US/10/080,505
? CURRENT FILING DATE: 2002-02-22
? PRIOR APPLICATION NUMBER: US 08/296,791
? PRIOR FILING DATE: 1994-10-25
? PRIOR APPLICATION NUMBER: US 09/839,996
? PRIOR FILING DATE: 2001-04-20
? NUMBER OF SEQ ID NOS: 58
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 1
? LENGTH: 4319
? TYPE: DNA
? ORGANISM: Haemophilus influenzae
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (60)..(4241)
? OTHER INFORMATION:
US-10-080-505-1

```

Best Local Similarity 100.0%; Pred. No. 0;
Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 3;

[illegible]


```

QY 3241 AAGAGAGAGTGTTCCTGATCCCTGCTGATCAAAAGCCTGTGGATTAAGAGCGAC 3300
DB 3241 AAGAGAGAGTGTTCCTGATCCCTGCTGATCAAAAGCCTGTGGATTAAGAGCGAC 3300
QY 3301 TTGAGGTATTTGATCCCAAGCAATCGGAAAAAGATGCTGCTCAAGAAAGCGG 3360
DB 3301 TTGAGGTATTTGATCCCAAGCAATCGGAAAAAGATGCTGCTCAAGAAAGCGG 3360
QY 3361 AAAAACAAGCAAAAGAAAGCTGATGACCGCTTATTCAAATAGGCTTATGAGAT 3420
DB 3361 AAAAACAAGCAAAAGAAAGCTTATGATGACCGCTTATTCAAATAGGCTTATGAGAT 3420
QY 3421 TATGCAACAGTAAATAGTATGCTTCTGTTCAGATGAATAGATGCTTTTGTG 3480
DB 3421 TATGCAACAGTAAATAGTATGCTTCTGTTCAGATGAATAGATGCTTTTGTG 3480
QY 3481 ATCAAGCAATCTGCGGTGAGCAAAATTCGACAGATTAAGACGTATGATTTG 3540
DB 3481 ATCAAGCAATCTGCGGTGAGCAAAATTCGACAGATTAAGACGTATGATTTG 3540
QY 3541 ATGCGTTCGCTGCTTATGACAGCAAAAGCACTTACATTTGGGGTGCAGAAAG 3600
DB 3541 ATGCGTTCGCTGCTTATGACAGCAAAAGCACTTACATTTGGGGTGCAGAAAG 3600
QY 3601 CCTTAGCTAAATGACCAATGGGGCAGTTTCTCGCATAGCCGTTGAGATTAACCTTG 3660
DB 3601 CCTTAGCTAAATGACCAATGGGGCAGTTTCTCGCATAGCCGTTGAGATTAACCTTG 3660
QY 3661 ATGAAACAGTAAATATCAAGCAATTAACGATGATGCTGGGTTTCCCAATTCAT 3720
DB 3661 ATGAAACAGTAAATATCAAGCAATTAACGATGATGCTGGGTTTCCCAATTCAT 3720
QY 3721 GGGGCGATTTACATTTGGTGTAAACGTGGGAACGGGAATCAGTGCAGTAAATGCGTG 3780
DB 3721 GGGGCGATTTACATTTGGTGTAAACGTGGGAACGGGAATCAGTGCAGTAAATGCGTG 3780
QY 3781 AAGAAACAACCCAAAAATTTATGAAAACGAAATTAATAGCGTAATGTAAGTATC 3840
DB 3781 AAGAAACAACCCAAAAATTTATGAAAACGAAATTAATAGCGTAATGTAAGTATC 3840
QY 3841 AGTTCGCTTAGGGCAATGGGCAATTCAGCTTATTTGGAGTAAATCCGATTTTATG 3900
DB 3841 AGTTCGCTTAGGGCAATGGGCAATTCAGCTTATTTGGAGTAAATCCGATTTTATG 3900
QY 3901 AACGTGAAATTTCAATCTGAGAGAGTGAAGTGAAGCCCTAGCGCTTCAATTATC 3960
DB 3901 AACGTGAAATTTCAATCTGAGAGAGTGAAGTGAAGCCCTAGCGCTTCAATTATC 3960
QY 3961 GCTATATGCTGCAATTCGAGTGAATTAATTAATCTCGACAGATTAATCAACGCTTA 4020
DB 3961 GCTATATGCTGCAATTCGAGTGAATTAATTAATCTCGACAGATTAATCAACGCTTA 4020
QY 4021 AACCTGATTTCTCGCATTAATGATGATTTCAAGAGCTTAACGTAACCAACGGTAA 4080
DB 4021 AACCTGATTTCTCGCATTAATGATGATTTCAAGAGCTTAACGTAACCAACGGTAA 4080
QY 4081 ATCTACAGGTGTGCAACAACATTTGAGAGTATTTGCAAAAAGAGAGTGGATTAAG 4140
DB 4081 ATCTACAGGTGTGCAACAACATTTGAGAGTATTTGCAAAAAGAGAGTGGATTAAG 4140
QY 4141 CAGAAATTTTACATTTCCAAATTTCCGCTTATTCGAAATCTCAAGGTTCAACCTG 4200
DB 4141 CAGAAATTTTACATTTCCAAATTTCCGCTTATTCGAAATCTCAAGGTTCAACCTG 4200
QY 4201 GCAAAACGCAAAATGGGGCGTAATTTGGGCTATGCTGTAATAATTAACATATTT 4260
DB 4201 GCAAAACGCAAAATGGGGCGTAATTTGGGCTATGCTGTAATAATTAACATATTT 4260
QY 4261 ATGCTTATTTGATTAACAAGGTGGGTGATGATGATCCCACTTTTATTCATTAAT 4319
DB 4261 ATGCTTATTTGATTAACAAGGTGGGTGATGATGATCCCACTTTTATTCATTAAT 4319

```

```

RESULT 4
PCT-US95-10661A-1
; Sequence 1, Application PCT/US9510661A
; GENERAL INFORMATION:
;   APPLICANT: Washington University, et al.
;   TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
;   NUMBER OF SEQUENCES: 9
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
;     STREET: 4 Embarcadero Center, Suite 3400
;     CITY: San Francisco
;     STATE: California
;     COUNTRY: United States
;     ZIP: 94111-4187
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: PCT/US95/10661A
;     FILING DATE: 16-AUG-1995
;   CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/296,791
;     FILING DATE: 25-AUG-1994
;   CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;     NAME: Treacathin, Richard F.
;     REGISTRATION NUMBER: 31,801
;     REFERENCE/DOCKET NUMBER: FP-59941/RFT
;     TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (415) 781-1989
;     TELEFAX: (415) 398-3249
;     TELEX: 916 277299
;   INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;     LENGTH: 4319 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: both
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 60..4241
;     PCT-US95-10661A-1
;
Query Match 100.0%; Score 4319; DB 5; Length 4319;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

361 GATATACAGATGTTGATTTTGTGTCAGAGGAAACAACCCGATCAACATCGTTTACTT 420
361 GATATACAGATGTTGATTTTGTGTCAGAGGAAACAACCCGATCAACATCGTTTACTT 420
421 ATTAAGATTGTAAGCAAAATTAACAACAAAAGATATTTTCACTTTATGAGACATTT 480
421 ATTAAGATTGTAAGCAAAATTAACAACAAAAGATATTTTCACTTTATGAGACATTT 480
481 ACCATATCCAGATTAATTAATGTTTACGAAAGGCGCTCCATATGATATGATCTGCA 540
481 ACCATATCCAGATTAATTAATGTTTACGAAAGGCGCTCCATATGATATGATCTGCA 540
541 ATATGATGGCAGTACTTATTCAGATTAAGCAAAATTAACAACCGATGTTGCTATCGCT 600
541 ATATGATGGCAGTACTTATTCAGATTAAGCAAAATTAACAACCGATGTTGCTATCGCT 600
601 CTGAGCGGCAAGTTTGGCGAAATGATCAAGCAAAAGCGACCAAGTTCCGGTGCATATC 660
601 CTGAGCGGCAAGTTTGGCGAAATGATCAAGCAAAAGCGACCAAGTTCCGGTGCATATC 660
661 ATATCTGACAGCTGGCAATTAACAATCAATCAAGCGGAGAGGATTAATGATATGATAT 720
661 ATATCTGACAGCTGGCAATTAACAATCAATCAAGCGGAGAGGATTAATGATATGATAT 720
721 TGGAGGCGCATGTTGCTTAAGCGGAGAAATATGCTCAATTAACGATTCAGGCTCAAGG 780
721 TGGAGGCGCATGTTGCTTAAGCGGAGAAATATGCTCAATTAACGATTCAGGCTCAAGG 780
781 GGGACAGTGGTCTCCGATGTTTATATGATGCTGAAAAAACAATAATGTTATATATG 840
781 GGGACAGTGGTCTCCGATGTTTATATGATGCTGAAAAAACAATAATGTTATATATG 840
841 GGAATATTAAGGGAAGGCAACCTTTGGAAGGAAAGAAATGCTTTCAATTTGATGCA 900
841 GGAATATTAAGGGAAGGCAACCTTTGGAAGGAAAGAAATGCTTTCAATTTGATGCA 900
901 AATCTTATTTGATGAATTTTGAAGAGATTTTACATATCACTTACACCCGACCTG 960
901 AATCTTATTTGATGAATTTTGAAGAGATTTTACATATCACTTACACCCGACCTG 960
961 GTAATGAGTGTACCAATTAATGATGATATGATGCTGAGGCGTCTATTAATCAAAAT 1020
961 GTAATGAGTGTACCAATTAATGATGATATGATGCTGAGGCGTCTATTAATCAAAAT 1020
1021 CAGGAATACCATCAAGAAATTAATTAAGTGAAGAAATAGATTAATGATTAATGATTA 1080
1021 CAGGAATACCATCAAGAAATTAATTAAGTGAAGAAATAGATTAATGATTAATGATTA 1080
1081 AGGATTAAGTGTATCTATGATATGATGAGGAAATTAATTAATGATTAATGATTAAT 1140
1081 AGGATTAAGTGTATCTATGATATGATGAGGAAATTAATTAATGATTAATGATTAAT 1140
1141 ATGAGGAAACGCTATATTTTATGATCAAAACCAAGATCACTTATCTGATCTGCA 1200
1141 ATGAGGAAACGCTATATTTTATGATCAAAACCAAGATCACTTATCTGATCTGCA 1200
1201 TTAACCAAGGCGGCGTCTTATTTGAGGAAATTAATTAAGTGAAGAAATAGATTAAT 1260
1201 TTAACCAAGGCGGCGTCTTATTTGAGGAAATTAATTAAGTGAAGAAATAGATTAAT 1260
1261 ACCAAACCTTGGCAAGAGCTGGCATATGATTAAGGAAATAGATTAATGATTAATGAT 1320
1261 ACCAAACCTTGGCAAGAGCTGGCATATGATTAAGGAAATAGATTAATGATTAATGAT 1320
1321 TTAATGCGGTGGAACATGATGATCTTCTTAATTTGATTAAGGAAATGATTAATGAT 1380
1321 TTAATGCGGTGGAACATGATGATCTTCTTAATTTGATTAAGGAAATGATTAATGAT 1380
1381 CCAAGGGAAGAAATTAAGGTTGATGATGAGGAGGATTAATGATTAATGATTAATGAT 1440
1381 CCAAGGGAAGAAATTAAGGTTGATGATGAGGAGGATTAATGATTAATGATTAATGAT 1440

1441 AGGACAGATCAAGGCAACAAACAGCCTTATGATTAATTTGCTTGGTTAGCGGAGAG 1500
1441 AGGACAGATCAAGGCAACAAACAGCCTTATGATTAATTTGCTTGGTTAGCGGAGAG 1500
1501 GGAATGTTCAATTAACGATGATTAACATTTGATACCGATTAATTTTATTTTGGCTTTC 1560
1501 GGAATGTTCAATTAACGATGATTAACATTTGATACCGATTAATTTTATTTTGGCTTTC 1560
1561 GTGGTGTGCTTATGATCTTAAACGCAATCACTTAACCTTTAAAGCTATCAAAATGAG 1620
1561 GTGGTGTGCTTATGATCTTAAACGCAATCACTTAACCTTTAAAGCTATCAAAATGAG 1620
1621 ACAGAGGCGCATATGATTAAGCAATTAATCAATCACTGACCGCTTAATGATTAATGAT 1680
1621 ACAGAGGCGCATATGATTAAGCAATTAATCAATCACTGACCGCTTAATGATTAATGAT 1680
1681 GGAAGGAAAGATTTGTTCTTCTTAATGATTAATTAATTAATTAATTAATTAATTAAT 1740
1681 GGAAGGAAAGATTTGTTCTTCTTAATGATTAATTAATTAATTAATTAATTAATTAAT 1740
1741 AAATGGCTTACAAAGTGGTTTGGCGAAACAGTAAATTAACATGATGGGCACTTA 1800
1741 AAATGGCTTACAAAGTGGTTTGGCGAAACAGTAAATTAACATGATGGGCACTTA 1800
1801 ACCTTATTTATTAACCAACCAAGATGCTACTTGTCTTCACTTCACTTCACTTCACTT 1860
1801 ACCTTATTTATTAACCAACCAAGATGCTACTTGTCTTCACTTCACTTCACTTCACTT 1860
1861 TAAAGGCGATTTTACCCAAACAAAGGTAATTTTTCAGCGGTGATGACCAACCGC 1920
1861 TAAAGGCGATTTTACCCAAACAAAGGTAATTTTTCAGCGGTGATGACCAACCGC 1920
1921 AAGCCTTACATCACTTAAATTAACGTTGATGAAATGAAAGGTAACCAAGGCGAA 1980
1921 AAGCCTTACATCACTTAAATTAACGTTGATGAAATGAAAGGTAACCAAGGCGAA 1980
1981 TTGTGTGGATCAAGATGATCAACCGTATCACTTAATTAAGCTTAATTAATTAATTAAT 2040
1981 TTGTGTGGATCAAGATGATCAACCGTATCACTTAATTAAGCTTAATTAATTAATTAAT 2040
2041 GGGAGAGTGGCGTGGTTCTGCGAATGTTTCTCAATGAGGAAATGATGAGCGTCAAG 2100
2041 GGGAGAGTGGCGTGGTTCTGCGAATGTTTCTCAATGAGGAAATGATGAGCGTCAAG 2100
2101 ATTAAGCAACGCACTTTGTTGTTGCAATCAACAAATCACTTTCACGCGT 2160
2101 ATTAAGCAACGCACTTTGTTGTTGCAATCAACAAATCACTTTCACGCGT 2160
2161 CAGATGGAACGATTAAGCACTTGTCAAAAGTGAATTAACCGATACAAAGTATTA 2220
2161 CAGATGGAACGATTAAGCACTTGTCAAAAGTGAATTAACCGATACAAAGTATTA 2220
2221 ATTCTTACCAAAACCAATCAATGAGCTGCTATTAATTAATGATTAATGATTAATGAT 2280
2221 ATTCTTACCAAAACCAATCAATGAGCTGCTATTAATTAATGATTAATGATTAATGAT 2280
2281 ATGTTAAGTGTGCAAAACCTTAATGCAATGCACTTAAACCAATCAACGCAATTA 2340
2281 ATGTTAAGTGTGCAAAACCTTAATGCAATGCACTTAAACCAATCAACGCAATTA 2340
2341 CATTAAGCAACATGCAACCAATGCAATGCAATGCACTTTCGCAATTAATGCAATGCA 2400
2341 CATTAAGCAACATGCAACCAATGCAATGCAATGCACTTTCGCAATTAATGCAATGCA 2400
2401 CGGTGATTAAGCAACCTTAATGCAATGCAATGCAATGCACTTTCGCAATTAATGCA 2460
2401 CGGTGATTAAGCAACCTTAATGCAATGCAATGCAATGCACTTTCGCAATTAATGCA 2460
2461 TAAAGGCGATTTTTCGCAACCAATTCAGGAGCAACAGGCAACAGTGAAGTGG 2520
2461 TAAAGGCGATTTTTCGCAACCAATTCAGGAGCAACAGGCAACAGTGAAGTGG 2520
2521 AAAATGCGATTTGCAACATGCTTACGATTAATGCAATTAATTAATTAATTAATTA 2580


```

Db 2521 AAAATGCACTGACAAATGCTACGACTACTCATTCGAAATTTTAACTAAATACA 2590
Qy 2581 GTAGATCAAGTTAAATCAAGTTTACAGTACGCTCAACAAATACGACGCGCTT 2640
Db 2581 GTAGATCAAGTTAAATCAAGTTTACAGTACGCTCAACAAATACGACGCGCTT 2640
Qy 2641 CATTAGAGACGAAACAAACGCAATCGGACAGACATCGTTTCAACATTTAGACATTA 2700
Db 2641 CATTAGAGACGAAACAAACGCAATCGGACAGACATCGTTTCAACATTTAGACATTA 2700
Qy 2701 ATGCTAATTTAGTGGGCAAGGCAATTCATTTCTTCTTCTTCTTCTTCTTCTTCTT 2760
Db 2701 ATGCTAATTTAGTGGGCAAGGCAATTCATTTCTTCTTCTTCTTCTTCTTCTTCTT 2760
Qy 2761 GCGATTAATTTAAATTTCAATACGCTGAGGAGGATTAATTTCTTCTTCTTCTTCTT 2820
Db 2761 GCGATTAATTTAAATTTCAATACGCTGAGGAGGATTAATTTCTTCTTCTTCTTCTT 2820
Qy 2821 CAGGCAAGAAACCCGAAACCTTGACATTTAATTTGTTGAAAGCAAGATTAATCAAC 2880
Db 2821 CAGGCAAGAAACCCGAAACCTTGACATTTAATTTGTTGAAAGCAAGATTAATCAAC 2880
Qy 2881 CGTTATCGAATAGCTCAATTTTCTTAAATAATCAAGATTAATCAAGATTAATCAAC 2940
Db 2881 CGTTATCGAATAGCTCAATTTTCTTAAATAATCAAGATTAATCAAGATTAATCAAC 2940
Qy 2941 GTTATTAATTTAGTGAAGATGATGGGCAATTCGCTGATTAATTTCTTCTTCTTCTTCTT 3000
Db 2941 GTTATTAATTTAGTGAAGATGATGGGCAATTCGCTGATTAATTTCTTCTTCTTCTTCTT 3000
Qy 3001 AATTGCAATGATTTAGTGAAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 3060
Db 3001 AATTGCAATGATTTAGTGAAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 3060
Qy 3061 TTGAACGCACTGCTAAACAAACAGGTAGACAAAGGTAGACAAAGGTAGACAAAGGTAG 3120
Db 3061 TTGAACGCACTGCTAAACAAACAGGTAGACAAAGGTAGACAAAGGTAGACAAAGGTAG 3120
Qy 3121 GAGCAGGCTTTCTGATACCTGCTGATCAAAAGCTGTTAAAGCTTGAAGCAAC 3180
Db 3121 GAGCAGGCTTTCTGATACCTGCTGATCAAAAGCTGTTAAAGCTTGAAGCAAC 3180
Qy 3181 AAGCTGAATGACTGCTGAACAAACAAAGGTAGACAAAGGTAGACAAAGGTAGACAA 3240
Db 3181 AAGCTGAATGACTGCTGAACAAACAAAGGTAGACAAAGGTAGACAAAGGTAGACAA 3240
Qy 3241 AAAAGAGAGTGTCTTCTGATCCCTGCTGATTAAGAGCTGTTGAAGCCGAC 3300
Db 3241 AAAAGAGAGTGTCTTCTGATCCCTGCTGATTAAGAGCTGTTGAAGCCGAC 3300
Qy 3301 TTGAGGTATTTAGTCCCAACAGCAATCGGAAAGATGCTGCTCAAGAAAGAGG 3360
Db 3301 TTGAGGTATTTAGTCCCAACAGCAATCGGAAAGATGCTGCTCAAGAAAGAGG 3360
Qy 3361 AAAAACAACCAACAAACAAACCTTGAATCAAGCCGTTATTAATGAGTATTAAGAT 3420
Db 3361 AAAAACAACCAACAAACAAACCTTGAATCAAGCCGTTATTAATGAGTATTAAGAT 3420
Qy 3421 TATCGCAAGTAAATAGTATGCTTCTGATCAAGATTAAGATTAAGATTAAGATTAAG 3480
Db 3421 TATCGCAAGTAAATAGTATGCTTCTGATCAAGATTAAGATTAAGATTAAGATTAAG 3480
Qy 3481 ATCAAGACAAATTCGCGTGTGACAAATTCGACAGAGATTAAGAGCTATGATTCG 3540
Db 3481 ATCAAGACAAATTCGCGTGTGACAAATTCGACAGAGATTAAGAGCTATGATTCG 3540
Qy 3541 ATGCGTTCCGCTTATCAAGACAGAAACAACTTACGCAATTTGGGTCGAAAGAG 3600
Db 3541 ATGCGTTCCGCTTATCAAGACAGAAACAACTTACGCAATTTGGGTCGAAAGAG 3600
Qy 3601 CTTAGCTAATGACAAATGAGGAGCTTTCTGATGAGCGTTGAGATTAATACCTTCTG 3660

```

```

Db 3601 CTTAGCTAATGACAAATGAGGAGCTTTCTGATGAGCGTTGAGATTAATACCTTCTG 3660
Qy 3661 ATGAAACAGGTTAAATCAAGCAATTAACGATGATGCGGCTTTTCCCAATTAATCAAT 3720
Db 3661 ATGAAACAGGTTAAATCAAGCAATTAACGATGATGCGGCTTTTCCCAATTAATCAAT 3720
Qy 3721 GGGGCAATTTACATTTGTTTAAACGTGGAGACGAGATCAATGAGATTAATTTAGCTTG 3780
Db 3721 GGGGCAATTTACATTTGTTTAAACGTGGAGACGAGATCAATGAGATTAATTTAGCTTG 3780
Qy 3781 AAGAACAAACCGGAAATTTCAATGAAAGAGATTAATTTGCGGTGAATGCAATGATTC 3840
Db 3781 AAGAACAAACCGGAAATTTCAATGAAAGAGATTAATTTGCGGTGAATGCAATGATTC 3840
Qy 3841 AGTTCCGTTTAAAGGCAATGAGGCAATTCAGCTTATTTTGAAGTAAATGCTATTTTATG 3900
Db 3841 AGTTCCGTTTAAAGGCAATGAGGCAATTCAGCTTATTTTGAAGTAAATGCTATTTTATG 3900
Qy 3901 AAGGTAAATTAATCAATCTGAGAAAGTGAAGTGAAGAAACGCTGACCTTGCATTTATC 3960
Db 3901 AAGGTAAATTAATCAATCTGAGAAAGTGAAGTGAAGAAACGCTGACCTTGCATTTATC 3960
Qy 3961 GCTATTAATGCTGCAATTCGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 4020
Db 3961 GCTATTAATGCTGCAATTCGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 4020
Qy 4021 AGCTTATTTCTTCTGATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 4080
Db 4021 AGCTTATTTCTTCTGATTAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 4080
Qy 4081 AATTCAGGCTGTTGCAACAACTTTGACGTTATTTGACAAAGAAAGATGAGATTAAG 4140
Db 4081 AATTCAGGCTGTTGCAACAACTTTGACGTTATTTGACAAAGAAAGATGAGATTAAG 4140
Qy 4141 CAGAAATTTCAATTTCAATTTTCTGCTTATTAATTTCAATTTCAATTTCAATTTCAATTT 4200
Db 4141 CAGAAATTTCAATTTCAATTTTCTGCTTATTAATTTCAATTTCAATTTCAATTTCAATTT 4200
Qy 4201 GCAACAGCAAAATGAGGAGGATTAATTTGAGCTATTTGATTTGATTTGATTTGATTTG 4260
Db 4201 GCAACAGCAAAATGAGGAGGATTAATTTGAGCTATTTGATTTGATTTGATTTGATTTG 4260
Qy 4261 ATGCTTATTTGATTAACAGGTGAGTCAATTCACCTTTTATTTCAATTAAT 4320
Db 4261 ATGCTTATTTGATTAACAGGTGAGTCAATTCACCTTTTATTTCAATTAAT 4320

```

```

RESULT 5
US-10-080-505-8
: Sequence 8, Application US/10080505
: Patent No. 6676948
: GENERAL INFORMATION:
: APPLICANT: St. Gene, Joseph W.
: TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
: FILE REFERENCE: A-59941-/RFT/DCF/DEH
: CURRENT FILING DATE: 2002-02-22
: PRIOR APPLICATION NUMBER: US 08/296,791
: PRIOR FILING DATE: 1994-10-25
: PRIOR APPLICATION NUMBER: US 09/839,996
: PRIOR FILING DATE: 2001-04-20
: NUMBER OF SEQ ID NOS: 58
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 8
: LENGTH: 4305
: TYPE: DNA
: ORGANISM: Haemophilus influenzae
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (1702)..(1702)
: OTHER INFORMATION: "a" at position 1702 can be any base.
: NAME/KEY: CDS
: LOCATION: (1)..(4305)

```

OTHER INFORMATION:

US-10-080-505-8

Query Match 73.6%; Score 3180.8; DB 4; Length 4305;
 Best Local Similarity 85.7%; Pred. No. 0;
 Matches 3704; Conservative 0; Mismatches 463; Indels 156; Gaps 9;

QY 60 ATGAAAAAACTGATTTCTGCTTAATTTTTTAAACCGCTGCATTTCAATTAAGATGTA 119
 Db 1 ATGAAAAAACTGATTTCTGCTTAATTTTTTAAACCGCTGCATTTCAATTAAGATGTA 60
 QY 120 TCGAACCGTGGCTGGTCACTTATTTGGATTGATTCATTAATTAATTTGATTTT 179
 Db 61 TCGAACCGTGGCTGGTCACTTATTTGGATTGATTCATTAATTAATTTGATTTT 120
 QY 180 GCGAGAAATTAAGGGAGTTCAAGTTGGGGCTCAAAATTAATTAAGTTATTAACAACAA 239
 Db 121 GCGAGAAATTAAGGGAGTTGCAAGTTGGGGCTCAAAATTAATTAAGTTATTAACAAGAA 180
 QY 240 GGGCAATTAGTTGGCAATCAATGACAAAAGCCCGATGATGATTTTCTGTAAGTCA 299
 Db 181 GGGCAATTAGTTGGCAATCAATGACAAAAGCCCGATGATGATTTTCTGTAAGTTC 240
 QY 300 CTTAAGCGCTGGCAGCTTGGTTGAATCAATATTTGAGCGTGGCAGATTAAGTA 359
 Db 241 AGAATGAGATTGCTGCTTGTAGTGGGATCAATATTTGAGTGGCAGATTAAGTA 300
 QY 360 GGATATZCAGATGTTGATTTGGTGCAGAGGGAAACAAACCCCGATCAATGCTTTACT 419
 Db 301 GGCTATZCAGATGTTGATTTGGTGCAGAGGGAAACAAACCCCGATCAATGCTTTACT 360
 QY 420 TATAAGTTGTAAACGAATTAATCAAAAGATTAATTAATTAATTAATTAATTAATTA 479
 Db 361 TATAAGTTGTAAACGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420
 QY 480 TACCATTAATCCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 539
 Db 421 TACCATTAATCCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
 QY 540 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 599
 Db 481 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
 QY 600 TCTGGACGCGATTTTGGCGAATGATCAAGAACAAAGCGACCAAGTTCGATAT 659
 Db 541 TCTGGACGCGATTTTGGCGAATGATCAAGAACAAAGCGACCAAGTTCGATAT 600
 QY 660 CATTATCTGACGCTGGCAATCAACATTCAGCTGGAGCAAGTATGATATTCGAT 719
 Db 601 AATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
 QY 720 TTGGAGGCGATTTTGGTAAAGGGGAAATTAATTAATTAATTAATTAATTAATTAAT 779
 Db 661 GTAGAGGCGATTTTGGTAAAGGGGAAATTAATTAATTAATTAATTAATTAATTAAT 720
 QY 780 GGGGACAGTGGTCTCGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 839
 Db 721 GGGGACAGTGGTCTCGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
 QY 840 GGGATTTACGGGAGGCAACCTTTTGAAGCAAGAAATGGCTTCAATGGTTCG 899
 Db 781 GGGATTTACGGGAGGCAACCTTTTGAAGCAAGAAATGGCTTCAATGGTTCG 840
 QY 900 AATCTTATTTGATGAATTTTGAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 959
 Db 841 AATCTTATTTGATGAATTTTGAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 900
 QY 960 GGTAAATGAGTGAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1019
 Db 901 GGTAAATGAGTGAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960
 QY 1226 TCGGAATACCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1079

Db 961 ATTGAAAAAC-----TCGTGAAATTAACCTTTACCGAACAAATTAATTAATTAATTAAT 1014
 QY 1080 AAGATAAAGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1139
 Db 1015 AAGATAAAGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1074
 QY 1140 AATGAGAAAGCTATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1199
 Db 1075 AATGAGAAAGCTATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1134
 QY 1200 AATTAACCAAGGCGGCTGATCTTATTTTGAAGGATTAATTAATTAATTAATTAATTAAT 1259
 Db 1135 AATGATCAAGGCTGAGCGGCTTGTATTTTGAAGGATTAATTAATTAATTAATTAATTAAT 1194
 QY 1260 AACCAACCTTGGCAAGAGCTGGCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1319
 Db 1195 AATGAAAGCTGAAGAGCTGGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1254
 QY 1320 GTAAATGGCTGGAACAATGATGATCTTTCTAAATTTGTTAAAGGAACTTGCACGTTCAA 1379
 Db 1255 GTAAATGGCTGGAACAATGATGATCTTTCTAAATTTGTTAAAGGAACTTGCACGTTCAA 1314
 QY 1380 GCGAAGGGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1439
 Db 1315 GCGAAGGGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1374
 QY 1440 CAGGACAGCATCAAGGCAACCAACAGCTTTAGTGAATTTGGCTGGTAGCGGACG 1499
 Db 1375 CAGGACAGCATCAAGGCAACCAACAGCTTTAGTGAATTTGGCTGGTAGCGGACG 1434
 QY 1500 GGGACGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1559
 Db 1435 GGGACGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1494
 QY 1560 CCGTGGTGGCTTGAATCTTAAGGCAATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1619
 Db 1495 CCGTGGTGGCTTGAATCTTAAGGCAATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1554
 QY 1620 GACGAGGGGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1679
 Db 1555 GACGAGGGGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1614
 QY 1680 GGGAAAGCAAGCAT--TGTTCTACCTTAATGAAATTAATTAATTAATTAATTAATTAATTAAT 1736
 Db 1615 GGGAAAGCAAGCAT--TGTTCTACCTTAATGAAATTAATTAATTAATTAATTAATTAATTAAT 1674
 QY 1737 AAGAAATTTGCTTAACAGCTTTGTTGGGAAACAGATTAATTAATTAATTAATTAATTAATTAAT 1796
 Db 1675 AAGAAATTTGCTTAACAGCTTTGTTGGGAAACAGATTAATTAATTAATTAATTAATTAATTAAT 1734
 QY 1797 TTAACCTTATTTTAACCAACCAAGAGATGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1856
 Db 1735 TTAACCTTATTTTAACCAACCAAGAGATGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1794
 QY 1857 AATTTAAAGGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1916
 Db 1795 AATTTAAAGGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1854
 QY 1917 CCGCACGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1976
 Db 1855 CCGCACGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1914
 QY 1937 GAAATTTGTGGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2036
 Db 1915 GAAATTTGTGGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1974
 QY 2037 AAGGCGGAAGTGGCTGCTTTCTGCAATGTTTCTTAATTAATTAATTAATTAATTAATTAATTAAT 2096
 Db 1975 AAGGCGGAAGTGGCTGCTTTCTGCAATGTTTCTTAATTAATTAATTAATTAATTAATTAATTAAT 2034
 QY 2097 AGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2156
 Db 2035 AGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2094

Db 4243 TCTCAAGTTCACTGCTAAACGAAATGCGGCTGGAATTCGCTATCGTTGG 4302
QY 4242 TAA 4244
Db 4303 TAA 4305

RESULT 6

US-09-557-884-1/C
Sequence 1, Application US/09557884
Patent No. 6506581

GENERAL INFORMATION:

APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884

FILING DATE: 25-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,102

FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PE186P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Query Match 69.9%; Score 3019.8; DB 4; Length 1830121;

Best Local Similarity 83.0%; Pred. No. 0;

Matches 3641; Conservative 2; Mismatches 650; Indels 93; Gaps 14;

QY 1 TCATATGCTTTAATA-TGATTTTAAACGAAATTAATCTTAATTAATTAACATT 59

Db 278587 TCATATGCTTTAACAAGTATTTTAAACGAAATTAATCTTAATTAATTAACATT 278528

QY 60 ATGAAAAAACTGATTTTCTTAATTTTAAACGCTGATTCATTAGAGATGTA 119

Db 278527 ATGAAAAAACTGATTTTCTTAATTTTAAACGCTGATTCATTAGAGATGTA 278468

QY 120 TCGAAGCTGGGCTGGTCACTTATTTGGATTGATTCACAAATTAATGATTTT 179

Db 278467 TCGAAGCTGGGCTGGTCACTTATTTGGATTGATTCACAAATTAATGATTTT 278408

QY 180 GCCGAGAAATTAAGGAGATTCACAGTTGGGGCTCAAAATATTAAGGTTATAGCAAA 239

Db 278407 GCCGAGAAATTAAGGAGATTCACAGTTGGGGCTCAAAATATTAAGGTTATAGCAAA 278348

QY 240 GGGCAATAGTTGGCAATCATGACAAAGCCCGATGATTTTCTGTAGTGTCA 299

Db 278347 GGGCAATAGTTGGCAATCATGACAAAGCCCGATGATTTTCTGTAGTGTGTG 278288

QY 300 GGTAAAGGCTGGGAGGCTTTGGTTGAAATATATATTTGGAGGCTGACATAAGTA 359

Db 278287 GGTAAAGGCTGGGAGGCTTTGGTTGAAATATATATTTGGAGGCTGACATAAGTA 278228

QY 360 GATATACAGATGTTGATTTTGGTGCAGAGGAAACACCCCGATCAATCGTTTACT 419

Db 278227 GATATACAGATGTTGATTTTGGTGCAGAGGAAACACCCCGATCAATCGTTTACT 278168

QY 420 TATAGATGTTAAACCAATTAATCTACAA---AAGATATTTTATCTCTTATAGAGAC 476

Db 278167 TATAGATGTTAAACCAATTAATCTACAA---AAGATATTTTATCTCTTATAGAGAC 278108

QY 477 GATTACCAATATCCAGCATTTACCTAAATTCGTTACAGAACGGCTCCATTTGATGACT 536

Db 278107 GATTACCAATATCCAGCATTTACCTAAATTCGTTACAGAACGGCTCCATTTGATGACT 278048

QY 537 TCGAATATGATGCGACTACTTATTCAGATAGAACAAATATCCAGAACTGTCTGATTC 596

Db 278047 TCGAATATGATGCGACTACTTATTCAGATAGAACAAATATCCAGAACTGTCTGATTC 277988

QY 597 GGCTCTGACGGCAGCTTTTGCGAAATGATCAAGACAAAGCG----- 639

Db 277987 GGCTCTGACGGCAGCTTTTGCGAAATGATCAAGACAAAGCG----- 277928

QY 640 -ACCAAGTCCCGTGCATATCTTGAACAGTGGCAATACACAAATCAACCGTGA 698

Db 277927 TATTATGCTCAGGTGCATATCTTGAACAGTGGCAATATCTTGAACAGTGGTGA 277868

QY 699 GCGATTAATGATATGTTGTTGGAGCGCATGTTGTTAAACGGGAGAAATAGTGTCA 758

Db 277867 GCGATTAATGATATGTTGTTGGAGCGCATGTTGTTAAACGGGAGAAATAGTGTCA 277808

QY 759 TTAACGATTCAGGCTCAAGGGGAGCAGTGGTCTCCAGTGTATTTATGATGCTGA 818

Db 277807 TTAACGATTCAGGCTCAAGGGGAGCAGTGGTCTCCAGTGTATTTATGATGCTGA 277748

QY 819 AAACAAAATGTTAATTAATGGAATTTACGGAGGCAACCTTTTGAAGCAAGAA 878

Db 277747 AAACAAAATGTTAATTAATGGAATTTACGGAGGCAACCTTTTGAAGCAAGAA 277688

QY 879 AATGGTTTCAATTTGTTGCAAACTTTATTTTAT-----GAATTTTC 923

Db 277687 AATGGTTTCAATTTGTTGCAAACTTTATTTTAT-----GAATTTTTC 277628

QY 924 GAAGAGATTTACATCATCATCTTACACCCGAGCTGTATGAGGTATACCAATTAGT 983

Db 277627 GAAGAGATTTACATCATCATCTTACACCCGAGCTGTATGAGGTATACCAATTAGT 277568

QY 984 GGAATGATTAATGCTCAGGGGCTATTAACGAAATACGGAATACATCGAAATTA 1043

Db 277567 GGAATGATTAATGCTCAGGGGCTATTAACGAAATACGGAATACATCGAAATTA 277508

QY 1844 AATAGCTTACGAATATGATTTACT-----TTGAAGAGAGAGATTAAGTTCAAT 1094

Db 277507 AATAGCTTACGAATATGATTTACT-----TTGAAGAGAGAGATTAAGTTCAAT 277448

QY 1835 AATCTAGATTAAGAGGACCTAATTTATTTCTCCAGTTTAAACATAGAGAAAGCTTA 1154

Db 277447 AATCTAGATTAAGAGGACCTAATTTATTTCTCCAGTTTAAACATAGAGAAAGCTTA 277388

QY 1155 TATTTATGATCAAAACAGGATCATTTCTTGCAATCTGACATTAACCAAGGGGCT 1214

Db 277387 TATTTATGATCAAAACAGGATCATTTCTTGCAATCTGACATTAACCAAGGGGCT 277328

QY 1215 GGGGCTTTATTTTGAAGGATTTTACATGATCTGCA---ATTCACCAAACTTGG 1271

Db 277327 GGGGCTTTATTTTGAAGGATTTTACATGATCTGCA---ATTCACCAAACTTGG 277268

QY 1272 CAGGAGCTGGCATATGTAAGTGAATAGCACCGTTTACTGGAAAGTAAATGCGGCT 1331

Db 277267 CAGAGTGGGCGTATCTATGAGCAAGATGCAACTGTTGAATGGAAGTTCCAAATCCT 277208
Qy 1332 GAACTAGTGCATCTTCTAAATTTGGTAAAGGAACATGCACTGCAAGCCAAAGGCAAA 1331
Db 277207 GAAATAGTGCATCTTCTAAATTTGGTAAAGGAACATGCACTGCAAGCCAAAG 277148
Qy 1392 AATTAAGTTCGATCAGCGTAGCGATGTAAGTCAATTTGGAGCAGCAGCAGAT 1451
Db 277147 AATTAAGTTCGATCAGCGTAGCGATGTAAGTCAATTTGGAGCAGCAGCAGAT 277088
Qy 1452 CAGGCAACAAACAGCCTTTAGTGAATTTGGCTTGTAGCGGAGAGGACTGTGCA 1511
Db 277087 GCGGGTCAAAACAGCCTTTAGTGAATTTGGCTTGTAGCGGAGAGGACTGTGCA 277028
Qy 1512 TTAAGCATGATTAAGCAATTTGATACCGATTAATTTGATTTGGCTTGTAGCGG 1571
Db 277027 TTAAGCATGATGATTAAGCAATTTGATACCGATTAATTTGATTTGGCTTGTAGCGG 276968
Qy 1572 TTAAGCATGATGATTAAGCAATTTGATACCGATTAATTTGATTTGGCTTGTAGCGG 1631
Db 276967 TTAAGCATGATGATTAAGCAATTTGATACCGATTAATTTGATTTGGCTTGTAGCGG 276908
Qy 1632 ATGATTTGGAACATTAATTAACATTAATTAATTAATTAATTAATTAATTAATTAAT 1691
Db 276907 ATGATTTGGAACATTAATTAACATTAATTAATTAATTAATTAATTAATTAAT 276848
Qy 1692 AT---TGTCTACCTTAATGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1748
Db 276847 ATTTACTGCTCATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276788
Qy 1749 TACAACGGTTGGTTGGCGAATCAGATTAATTAATTAATTAATTAATTAATTAATTAAT 1808
Db 276787 TACAACGGTTGGTTGGCGAATCAGATTAATTAATTAATTAATTAATTAATTAAT 276728
Qy 1809 TTTAAACCAACCAACAAATTTGATTTGCTTCACTTCACTTCACTTCACTTCACTTCACT 1868
Db 276727 TTTAAACCAACCAACAAATTTGATTTGCTTCACTTCACTTCACTTCACTTCACT 276668
Qy 1869 GATATTCCCAACCAACAAATTTGATTTGCTTCACTTCACTTCACTTCACTTCACTTCACT 1928
Db 276667 GATATTCCCAACCAACAAATTTGATTTGCTTCACTTCACTTCACTTCACTTCACT 276608
Qy 1929 AATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1988
Db 276607 AATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276548
Qy 1989 GATCAACATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2048
Db 276547 GATCAACATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276488
Qy 2049 GCGGTGTTCTCGCAATTTGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 2108
Db 276487 GCGGTGTTCTCGCAATTTGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 276428
Qy 2109 AATGCCACATTTGTTGTGTCGAATTAATTAATTAATTAATTAATTAATTAATTAAT 2168
Db 276427 AATGCCACATTTGTTGTGTCGAATTAATTAATTAATTAATTAATTAATTAATTAAT 276368
Qy 2169 AAGAGATTAACGATTTGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 2228
Db 276367 AAGAGATTAACGATTTGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 276308
Qy 2229 CCAAAAAACAAATCAATGCTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2288
Db 276307 CCAAAAAACAAATCAATGCTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276248
Qy 2288 GGTTCAGAAAACCTTAATGCAATGCACTTTAACAATTAATTAATTAATTAATTAATTAAT 2348
Db 276247 GGTTCAGAAAACCTTAATGCAATGCACTTTAACAATTAATTAATTAATTAATTAATTAAT 276188
Qy 2349 AACAAAGCAGCCAAATTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2408

Db 276187 AACAAAGCAGCCAAATTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 276128
Qy 2409 AATGCAAACTGAAACGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2468
Db 276127 AATGCAAACTGAAACGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276068
Qy 2469 AGCAATTTTGGCAACAAATTTGAGGAGACAAAGGACAAACAGTACGTTGAAATGCG 2528
Db 276067 AGCAATTTTGGCAACAAATTTGAGGAGACAAAGGACAAACAGTACGTTGAAATGCG 276008
Qy 2529 ACTTGAACATGCTTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2588
Db 276007 ACTTGAACATGCTTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275948
Qy 2589 ACCTTAATTTGCTTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2645
Db 275947 ACCTTAATTTGCTTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275888
Qy 2646 GAGACGGAACCAACGCAATGCGAGACATGCTTGAACATTTGACAGTAAATGCT 2705
Db 275887 GAGACGGAACCAACGCAATGCGAGACATGCTTGAACATTTGACAGTAAATGCT 275828
Qy 2706 AATTTGATGCGCAAGCAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2765
Db 275827 AATTTGATGCGCAAGCAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275768
Qy 2766 AATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2825
Db 275767 AATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275708
Qy 2826 AAGAACCCGAAACCTTTGACATTAATTTGCTTGAAGAAAGATTAATTAATTAATTAAT 2885
Db 275707 AAGAACCCGAAACCTTTGACATTAATTTGCTTGAAGAAAGATTAATTAATTAATTAAT 275648
Qy 2886 TCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2945
Db 275647 TCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275588
Qy 2946 AATTTGTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3005
Db 275587 AATTTGTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275528
Qy 3006 CAAATGATTTAGTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3065
Db 275527 CAAATGATTTAGTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 275468
Qy 3066 CCGATGCTTAACCAACAAAGTGAAGCAAAATGCTGCAAGAAAGCAAGCAAGCAAGCAAG 3125
Db 275467 CCGATGCTTAACCAACAAAGTGAAGCAAAATGCTGCAAGAAAGCAAGCAAGCAAG 275417
Qy 3126 GCGTTTCTGATTAACCTGCTGATCAAAAGCTGTTAAAGCATTTGAAGCAAGCAAGCT 3185
Db 275416 GCGTTTCTGATTAACCTGCTGATCAAAAGCTGTTAAAGCATTTGAAGCAAGCAAGCT 275357
Qy 3186 GAACTGATGCTGAACCAACAAAGTGAAGCAAAATGCTGCAAGAAAGCAAGCAAGCAAG 3245
Db 275356 GAACTGATGCTGAACCAACAAAGTGAAGCAAAATGCTGCAAGAAAGCAAGCAAGCAAG 275297
Qy 3246 -----GAGTGTCTTCTGATTCCTGCTGATTAAGTCTGCTGCTGCTGCTGCTGCTGCT 3293
Db 275296 GAGTGTGATGAGGCTTCTGATTCCTGCTGATTAAGTCTGCTGCTGCTGCTGCTGCTGCT 275237
Qy 3294 GCGGACCTTGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3353
Db 275236 GCGGACCTTGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275177
Qy 3354 GAAAGGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 3413
Db 275176 G-----AAGAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 275126
Qy 3414 TCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3473
Db 275125 TCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 275066

OY	374	TTTGTGATCAAGACCATCTGCAGTGTCGAATAATCGCACAAGATAAAAAGAGCTAT	3533
Dd	275065	TTTTGATTATTAAGCACAAATCTGCCCTGTGGCAAATATGTCCACAGATTAAGAAGCGCTAT	275006
OY	3534	GATTCTGATGTCGTTCCGTGCTTATCAGCAGCAAGAAAAAGAACTTAATGCTAATTTGGGGTG	3593
Dd	275005	GATTCTGATGTCGTTCCGTGCTTAT---CAGCAGAAAAAGAACTTAATGCTAATTTGGGGTG	274949
OY	3594	CAAAAAGCCTTAGCTAATGACGAATTTGGGGCAGTTTCTGCATAGCCGTTCAATAT	3633
Dd	274948	CAAAAAGCCTTAGCTAATGACGAATTTGGGGCAGTTTCTGCATAGCCGTTCAATAT	274889
OY	3654	ACCTTATGATGAACAGTTTAAAAATCAGCGACATTACAGATGAATGTGGGGTTTGCCAA	3713
Dd	274888	ACCTTATGATGAACAGTTTAAAAATCAGCGACATTACAGATGAATGTGGGGTTTGCCAA	274829
OY	3714	TATCAATGGGGGAGTTTACATTTGGTGTAAACGTGGGAACGGGAATCAGTGCAGATAA	3773
Dd	274828	TATCAATGGGGGAGTTTACATTTGGTGTAAATGTGGGTGGGAATTAATGTGCAGATAA	274769
OY	3774	ATGCGTGAAGAACAAAGCCGAAAAATTCATGAAAGCGTAAATTAATGCGGTGAATGCA	3833
Dd	274768	ATGCGTGAAGAACAAAGCCGAAAAATTCATGAAAGCGTAAATTAATGCAAGATGCA	274709
OY	3834	AGTATCAGTGCAGTTTATGGGCAATTTGGGCAATTCAGCCTTATTTTGGAGTTATGCTAT	3893
Dd	274708	AGTATCAGTGCAGTTTATGGGCAATTTGGGCAATTCAGCCTTATTTTGGAGTTATGCTAT	274649
OY	3894	TTTATTAACGTGAATAATTTATCAATCTGAGAGTGAAGTGAAGAAAGCGCTACGCTTCCA	3953
Dd	274648	TTTATTAACGTGAATAATTTATCAATCTGAGAGTGAAGTGAAGTGAAGAAAGCGACCTTCCA	274589
OY	3954	TTTATTCGCTAATTAATGCTGGCATTTGAGTTGATTAATCACTTACTCGACAGATAATTC	4013
Dd	274588	TTTATTCGCTAATTAATGCTGGCATTTGAGTTGATTAATCACTTACTCGACAGATAATTC	274529
OY	4014	AGCGTAAAGCCTTATTTCTTCGTCATTAATGTGTAAATGTTTCAAAAGCGTAACGTAAACCAACC	4073
Dd	274528	AGCGTAAAGCCTTATTTCTTCGTCATTAATGTGTAAATGTTTCAAAAGCGTAACGTAAACCAACC	274469
OY	4074	AACGGTAATTCACAGGTGTGTGCACAACACATTTGAGCGTTATTTGGCAAAAAGAAAGTGGGA	4133
Dd	274468	ACTGTAAATTAAGCACAGATTTGTGCACAACATCATTTGGGGCTTAATTTGCAAAAAGAAAGTGGGA	274409
OY	4134	TTTAAGGCGAAATTTTATCATTTCCAAATTTCCGCTTTTATCTCAAAATCTCAAGGTTCA	4193
Dd	274408	TTTAAGGCGAAATTTTATCATTTCCAAATTTCTTCGCTTTTATTTCAAAATCTCAAGGTTG	274349
OY	4194	CATCTGGGCAACAGCAAAAATGTGGGCGTGAATTTGGGCTATCGTGGTAAATTAACAACA	4253
Dd	274348	CATCTGGGCAACAGCAAAAATGTGGGCGTGAATTTGGGCTATCGTGGTAAATTAACAACA	274289
OY	4254	TAAATTTATCGTTTATGATTAACAAAGTGGGTCAATGATCAATCCACTTTTATATTC	4313
Dd	274288	TAAATTTATCGTTTATGATTAACAAAGTGGG-----GAGATCCCACTTTTATATTC	274234
OY	4314	AATAAT 4319	
Dd	274233	AATAAT 274228	
 RESULT 7 US-09-643-990A-1/C ; Sequence 1, Application US/09643990A ; Patent No. 6528289 ; GENERAL INFORMATION: ; APPLICANT: Robert D. Fleischmann ; Mark D. Adams ; Owen White ; Hamilton O. Smith ; J. Craig Venter ; TITLE OF INVENTION: The Nucleotide sequence of			

1		NUMBER OF SEQUENCES: 1	
2		CORRESPONDENCE ADDRESS:	
3		ADDRESSEE: Human Genome Sciences, Inc.	
4		STREET: 9410 Key West Avenue	
5		CITY: Rockville,	
6		STATE: MD	
7		COUNTRY: USA	
8		ZIP: 20850	
9		COMPUTER READABLE FORM:	
10		MEDIUM TYPE: 3 1/2 inch diskette	
11		COMPUTER: Dell Pentium	
12		OPERATING SYSTEM: MS DOS v6.22	
13		SOFTWARE: ASCII Text	
14		CURRENT APPLICATION DATA:	
15		APPLICATION NUMBER: US/09/643,990A	
16		FILING DATE: 23-Aug-2000	
17		CLASSIFICATION: <Unknown>	
18		PRIOR APPLICATION DATA:	
19		APPLICATION NUMBER: 08/487,429	
20		FILING DATE: 1995-06-07	
21		APPLICATION NUMBER: 08/426,787	
22		FILING DATE: 1993-04-21	
23		ATTORNEY/AGENT INFORMATION:	
24		NAME: Kenley K. Hoover	
25		REGISTRATION NUMBER: 40,302	
26		REFERENCE/DOCKET NUMBER: PB186P1C1	
27		TELECOMMUNICATION INFORMATION:	
28		TELEPHONE: 301-610-5790	
29		TELEFAX: 310-309-8439	
30		INFORMATION FOR SEQ ID NO: 1:	
31		SEQUENCE CHARACTERISTICS:	
32		LENGTH: 1830121 base pairs	
33		TYPE: nucleic acid	
34		STRANDEDNESS: double	
35		TOPOLOGY: linear	
36		SEQUENCE DESCRIPTION: SEQ ID NO: 1:	
37		US-09-643-990A-1	
38		Query Match	69.3%; Score 3019.8; DB 4; Length 1830121;
39		Best Local Similarity	83.0%; Pred. No. 0;
40		Matches 3641; Conservative	2; Mismatches 650; Indels 93; Gaps 14;
41	QY	1 TCQAATAGTCGTTAACTA-GTATTTTTTAATACAAAAATTTCTTAATTAATAACAATT	59
42	DB	278587 TCAATAGTCGTTAAACCAACGTATTTTAAACGAAAATTAATTATTAATAAACACT	278528
43	QY	60 ATGAAAAAACGTATTTGGCTTAATTTTAAACCGCTGCATTTGATTAGAGGAATAGTA	119
44	DB	278527 ATGAAAAAACGTATTTGGCTTAATTTTAAACCGCTTGCTTTATTTAGCGCTTGCTTTATTTAGCATTA	278468
45	QY	120 TCGAAGCGCTGGCGTCACACTTATTTTGGGATTTGATTACCAATATTAATCGTAGATTTT	179
46	DB	278467 TCAAGAAGCTGGCGAGGCTCATCTATTTTGGGATTTGATTACCAATATTAATCGTAGATTTT	278408
47	QY	180 GCCGAAATTAAGGAGAGTTCACAGTGGGGCTCAAAATATTAAGCTTTTAAACAAACA	239
48	DB	278487 GCCGAAATTAAGGAGAGTTCACAGTGGGGCTCAAAATATTAAGGCTTTTAAACAAACA	278348
49	QY	240 GGCGAATTTAGTTGGCACACATATGACAAACCCCGCATGATTTGATTTTTCGTAGTGTCA	299
50	DB	278347 GGCGAATTTAGTTGGCACACATATGACAAACCCCGCATGATTTGATTTTTCGTAGTGTGTG	278288
51	QY	300 CGTAAAGCGGTGCGACGCCCTTGTTGAAAATCATATATTGTGAGCGTGGCACTAAAGTA	359
52	DB	278287 CGTAAAGCGGTGCGACGCCCTTGTTGAAAATCATATATTGTGAGCGTGGCACTAAAGTA	278228
53	QY	360 GGAATATACGATGTTGATTTTGGTGGACAGAGGAAACACCAGATCAACATCGTTTACT	419
54	DB	278227 GGAATATACGATGTTGATTTTGGTGGACAGAGGAAACACCAGATCAACATCGTTTACT	278168
55	QY	420 TAATAGATTGAAAACGAATATTAACA--AAAGATTAATTATTCATTCCTTATTAAGGAC	476

Db 278161 TTTCAATTGTAAAAAGAAATTAATATCTCAAGCTTGGAGAGAAAGCAATCCTTATGATGGA 278108

Oy 477 GATTACCAATATTCACAGATTACATAAATTGGTTACAGAACGGCTCCAAATTGATGACT 536

Db 278107 GATTATCATATGCTCTGTTATCTACAAATTTGTAATCTGAGCTGAACTGTGGGTATGACA 278048

Oy 537 TCGAATATGAAATGGCAGATACCTATATTCAGATGAAACAAATATCCGAAACGTGTGCTATC 596

Db 278047 ACAAATATGGATGGAAAAAGTATATGCTGATGAGAGAACTATCTCGACCGGTGATATAT 277988

Oy 597 GGCCTCTGACGGCAGCTTTTGCGAAATGATCAAGACAAAGCG----- 639

Db 277987 GGCCTCAGAGATGATGATTTGGCGTAAGATTAAGATGAAGAAACCAATGTGATGATTCA 277928

Oy 640 -ACCAATGTCGGGTGCATATCATTTCTGACAGCTGGCAATPACCAATACGGTGG 698

Db 277927 TATTATCTCAGAGCAATATGCTATCTCTCTCAGAGAAATGCCATCTACGAGTGG 277868

Oy 699 GCAGATATGGAATATCGTATTTGGAGGCGGCTTGTGTAAAGCGGAGAAATTTGGTCCA 758

Db 277867 AATGTATATGTGACGTCAATCTTATGTGTATATGATGATGCTTAACTATATGTGTCA 277808

Oy 759 TTACCGATGCAAGCTCAAAAGGCGCAGATGCTTCCGATGTTTATTTATATGACTGAA 818

Db 277807 TTACCAACGGGTGTGTTCTAAAGCGATAGCGGTCCGCAATGTTATTTATGATGGAAG 277748

Oy 8-9 AACAATAATATGTTAAATTATAGGAATATTAGGGAAAGCAACCTTTGAAAGGCAAGAA 878

Db 277747 AAGCAACATGCTTATATATGCTGATATCAAACTGGGCATCTTTTTCGAGAGAGT 277688

Oy 879 AATGGTTTCAATGTTGTCGAATCTTATTTGAT-----GAAATTTTC 923

Db 277687 AATGGTTTCAATGTTTACGTGAGAAAGGTTTATATAGAGTTCTTGGGTGATACC 277628

Oy 924 GAAAGATTTTCAATCAATCACTTTACACCGAGCTGTATATGAGATGACAAATTGCT 983

Db 277627 CCTGATGTTTTCACGCTATATTTCCCAATAAATGACATTAATCTTTGTATCAAT 277568

Oy 984 GGAATATATATGTCACAGGGTCTATACATCAGAAATCAGGAATATCCATCAGAACTAA 1043

Db 277567 AATCATGTGACGTAATTAATTAATCTTAACTTACACTGTGATGAAGATGCTCTAAAGCAAA 277508

Oy 1044 ATTACGTTGACAAATATGATTACT-----TGAAAGAGAGATTAAGTTCAAT 1094

Db 277507 TCAAGATGAGAACTGTGAAGTATTAATCATCATGTTAAATCAAAACCGCTAAAGACAT 277448

Oy 1095 AATCCATGATATGACGACCTAATTAATTTCTCAGCTTAAACANTGAGAAAGCTA 1154

Db 277447 GTTAAACGACGACGCGTATATATTAATTAACAACAAGATGAAATGAAAAATATTT 277388

Oy 1155 TATTTATGATCAAAAACAAGANTCATTAATCTGTGACATGACATTAAACAAGAGGCG 1214

Db 277387 TACTTGCGACCAAGGAAAAAGAACTTTAACAATGAAATATATTAATCAAGTGTCT 277328

Oy 1215 GGTGCTTATTTTGAAGGTAATTTACATATTCGA--ATTCAACAATCTTG 1271

Db 277327 GGTGATATTAATCTTGAAGGTAAATTTTGTATAAAGGCAGACAAATATATATACTGG 277268

Oy 1272 CAAGAGATGAGCATCATGTAATGTAATATGACGCTTACTTGTGAAAGTAATGGCGTG 1331

Db 277267 CAAGTGCAGGGTATCATATGACAAAGTGCACGTGTGAATGSAATGTTCACAATCT 277208

Oy 1332 GAAATGATGCACTTTCTAAATTTGTATTAAGAAATCTGCAAGCTTCAAGCAAGGGAA 1391

Db 277207 GAAATGATGCTTTATCTAAATTTGTATATAGCACTTTATATGATCATGTTATAGGAAG 277148

Oy 1392 AATAAAGTTCATAGCGTAGCGCATGTGTAATCATTTTGTAGACGACGSCAGACAT 1451

Db 277147 AATTGSAAGTTTAATGTGGGTAAACGGCAATCATCTTGTGATCAACAGAGATGA 277088

Oy 1452 CAAGCAACAAGCTTATGTAATTTGGCTTGGTTAGCGCAGAGSACTGTTCA 1511

Dd	277087	GGGGGTTGAAAAACAAGCTTTCGAAAGAAAGTGGCACTGTGTAAGGGGTGCAAGCAACAGTTTCAA	277028
Qy	1512	TTAAGCATGATTAACAAATTTGATATACCGATTAATTTATTTTCGCTTGGCTGGTGGTGGC	1571
Dd	277027	TTAAATTTGTAACAGATCAAGTTGGATCCCTAACAAATCATATTTTCGATTTTCGGTGGTGGC	276968
Qy	1572	TTAATCTTTAAGGGGCAATCTAATACCTTTAAACGTATACGTAACAAATTAGGACGAGGGGGCA	1631
Dd	276967	TTAATCTTTAAGGGGCAATCTAATACCTTTAAACGTATACGTAACAAATTAGGACGAGGGGGC	276908
Qy	1632	ATGATTTGTGAACATAATATACACTCAAGCCGCTAAATGTCACTATTACTGAGAACGAAAGC	1691
Dd	276907	ATGATTTGTGAACATAATATACACTCAAGCCGCTAAATGTCACTATTACTGAGAACGAAAGT	276848
Qy	1692	AT--TGTCTACTTAATGTAATTAATTTAATTAACCTTGATTAACGAAAAGAAATGGC	1748
Dd	276847	ATTTCTGCTCCATCTTAATTAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTGCT	276788
Qy	1749	TACAGCGTTGGTTGGGCAACAGATTAATAATTTAAACAAATGCGGCGCTTAATACCTTAAT	1808
Dd	276787	TACAGCGTTGGTTGGGCAACAGATTAATAATTTAAACAAATGCGGCGCTTAATACCTTAAT	276728
Qy	1809	TATTAACCAACACAGAGATGTAATTTGCTACTTCAAGTGGTAACAAATTTAAAAAGC	1868
Dd	276727	TATTAACCAACACAGAGATGTAATTTGCTACTTCAAGTGGTAACAAATTTAAAAAGC	276688
Qy	1869	GATTTATACCAACAAAAGGTAAATTTTTCAGGGGTAGACGACACGACCGCTTAC	1928
Dd	276687	GATTTATACCAACAAAAGGTAAATTTTTCAGGGGTAGACGACACGACCGCGCTTAC	276608
Qy	1929	AATCATTTAATTAACGTTGGTACAGAAATGAGAGATACACAAAGCGAAATTTGGTG	1988
Dd	276607	AATCATTTAATTAACGTTGGTACAGAAATGAGAGATACCAACAGGCGAAATTTGGTG	276548
Qy	1989	GATCAGATTTGATACACGTAATTTTAAAGTGAACCTTCCAAATTTAAAGGGGAAAT	2048
Dd	276547	GATTTACATTTGATACACGTAATTTTAAAGTGAACCTTCCAAATTTAAAGGGGAAAT	276488
Qy	2049	GGGGTGGTTTCCGCAATGTTCTTCAATTAAGGCAAAATTTGACAGTCAACAATTAATGCA	2108
Dd	276487	GGGGTGGTTTCCGCAATGTTCTTCAATTAAGGCAAAATTTGACAGTCAACAATTAATGCA	276428
Qy	2109	AATGCCAATTTGGTGTGTGTCGCAATACACAAATATCAATTTGACGCGCTTCAATTTG	2168
Dd	276427	AATGCCAATTTGGTGTGTGTGTCGCAATATCAAAATATCAATTTGACGCGCTTCAATTTG	276368
Qy	2169	ACAGATTTAGACCTGTGCAAAATGGAATTTAACCGATACAAAGTATTAATCTTATA	2228
Dd	276367	ACAGATTTAGACCTGTGCAAAATGGAATTTAACCGATACAAAGTATTAATCTTATA	276308
Qy	2229	CGAAAAACCAATTCGCTCTATTAATTTAATCTGATATGCAAGCGGAATGTTAA	2288
Dd	276307	CGCAATACCAATTTAATGCTCTATTAATTTAATCTGATATGCAAGGAATTTAT	276248
Qy	2289	GGTTTGCAAAATTTAATGCAATGTCACTTTAAACAAATGCAAGCAATTTCACTTAAGC	2348
Dd	276247	GGTTTGCAAAATTTAATGCAATGTCACTTTAATTAATGCAAGCAATTTCACTTTAAGC	276188
Qy	2349	AACAATGCCAACCAATTTAGCAATATGCACTTCCGACAAATTAACGTGACGCTGAT	2408
Dd	276187	AACAATGCCAACCAATTTAGCAATATGCACTTCCGACAAATTTAATGCAAGCAATTTAAT	276128
Qy	2409	AATGCAAACTTGAACGTAATGTGCACTTAACGATTTCACTCAATTTCTTTAAAAAAC	2468
Dd	276127	AATGCCAAGTTAAAGGCAATGTGCACTTAACGATTTCTGCTCAATTTCTTTAAAAAAC	276068
Qy	2469	AGGCATTTTTCACAAATTTAGGGGACAAAGGACAAACAGTACGTTGGAAAATGCG	2528
Dd	276067	AGGCATTTTTCACAAATTTAGGGGACAAAGGACAAACAGTACGTTGGAAAATGCG	276008
Qy	2529	ACTTGCACATGCTTACGCAATACATTCATTCGAAATTTAAGCTTAATTAACATACATACATC	2588
Dd	276007	ACTTGCACATGCTTACGCAATACATTCATTCGAAATTTAAGCTTAATTAATTAATGCTTAT	275948

QY	3654	ACCTTTGATGAACAGGTTAAATAATTCACGGGACATTACCATGATGTCGGGTTTGCCCAA	3713
Db	274889	ACCTTTGATGAACAGGTTAAATAATTCACGGGACATTACCATGATGTCGGGTTTGCCCAA	274822
QY	3714	TATCATATGCGGCGATTTCACATTTGGTGTAAACGTGGGAAACGGGATCATGTCGAGTAAA	3773
Db	274828	TATCATATGCGGCGATTTCACATTTGGTGTAAACGTGGGAAATAGTCGAGTAAA	274769
QY	3774	ATGCTGTGAAGAAACAAGCGCCAAAATAATCATCTGAAAAGCATAAATATATGCGGTAAATGCA	3833
Db	274766	ATGCTGTGAAGAAACAAGCGCCAAAATAATCATCTGAAAAGCATAAATATATGAGAGTAATGCA	274705
QY	3834	AGTATCATGTTCCGTTTATAGGCGCAATTTGGCACTTTCAGGCTTAACTTTTGGAGTTAATCGCTAT	3893
Db	274708	AGTATCATGTTCCGTTTATAGGCGCAATTTGGGTAATTCAGGCTTATTTGGGTTGTTAATCATAT	274644
QY	3894	TTTATGTAACGTGAAATAATTCATTTGAGAGAGTGAAGTGAATAAGGCTTACGCTTGCA	3953
Db	274648	TTTATGTAACGTGAAATAATTCATTTGAGAGAGTGAAGTGAATAAGGCTTACGCTTGCA	274589
QY	3954	TTTATATGCTATATATGCTGGCATTGTGAGTGTATATCATATTACTCCGACAGATATATTC	4013
Db	274588	TTTATATGCTATATATGCTGGCATTGTGAGTGTATATCATATTACTCCGACAGATATATTC	274522
QY	4014	AGCGTTAAGCTTATTTCTTCTGCAATTATGTTGATGTTTCAAACGCTTAAAGTCAAAACC	4073
Db	274528	AGCGTTAAGCTTATTTCTTCTTCTGCAATTATGTTGATGTTTCAAACGCTTAAAGTCAAAACC	274469
QY	4074	ACGGTAAATCTCAGGTTGGTGGCAACAACATTTGGAGGTTATTTGGCAAAAAGAAAGTGGGA	4133
Db	274468	ACTGTAAATCTCAGGTTGGTGGCAACAACATTTGGAGGTTATTTGGCAAAAAGAAAGTGGGA	274405
QY	4134	TTAAAGGCAAAATTTTACATTTCCAAATTTTCGCTTTATCTCAAAATCTCAAGTTCA	4193
Db	274408	TTAAAGGCAAAATTTTACATTTCCAAATTTTCGCTTTATCTCAAAATCTCAAGTTCA	274345
QY	4194	CAACTCGGCAAAACGCAAAATGTGGGCGTGAATTTGGGCTATTCGTTGTTAAATATCAACA	4253
Db	274348	CAACTCGGCAAAACGCAAAATGTGGGCGTGAATTTGGGCGTATTCGTTGTTAAATATCAACA	274289
QY	4254	TAAATTTATCGTTTATTTGATTAACAAGGTGGTCAATGATGATATCCACCTTTTATTTCC	4313
Db	274288	TAAATTTATCGTTTATTTGATTAACAAGGTGGG-----GCAATATCCACCTTTTATTTTC	274233
QY	4314	AATTAAT 4319	
Db	274233	AATTAAT 274228	
RESULT 8			
US-ID-080-505-12			
Sequence 12, Application US/1008305			
Patent No. 6676948			
GENERAL INFORMATION:			
APPLICANT: St. Geme, Joseph M.			
TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS			
FILE REFERENCE: A-59941-1/RET/7CP/DBR			
CURRENT FILING DATE: 2002-02-22			
PRIOR APPLICATION NUMBER: US 08/296,791			
PRIOR FILING DATE: 1994-10-25			
PRIOR APPLICATION NUMBER: US 09/839,996			
PRIOR FILING DATE: 2001-04-20			
NUMBER OF SEQ ID NOS: 58			
SOFTWARE: PatentIn version 3.1			
SEQ ID NO 12			
LENGTH: 5245			
TYPE: DNA			
ORGANISM: Haemophilus influenzae			
FEATURE:			
NAME/KEY: CDS			
LOCATION: (430)..(4740)			
OTHER INFORMATION:			

US-10-080-505-12

Query Match 69.2%; Score 2990.4; DB 4; Length 5245;
 Best Local Similarity 82.0%; Pred. No. 0;
 Matches 3660; Conservative 0; Mismatches 636; Indels 166; Gaps 12;

1 TCATAGTCTTTAACTA-GTATTTTAAACGAAAATTACTTAATTAATTAACNTT 59
 379 TCATAGTCTTTAACTA-GTATTTTAAACGAAAATTACTTAATTAATTAACNTT 429
 60 ATGAAAAAACTGTAATTCCTTAATTTTAAACGCTTCATTTCAATTAAGGATAGA 119
 430 ATGAAAAAACTGTAATTCCTTAATTTTAAACGCTTCATTTCAATTAAGGATAGA 489
 120 TCGCAAGCTGGGCTGGTCA-CACTATTTTGGGATTTGATTAACATTAATTCGATATT 179
 490 TCGCAAGCTGGGCTGGTCA-CACTATTTTGGGATTTGATTAACATTAATTCGATATT 549
 180 GCCGAGATTAAGGAGGATTCACAGTTGGGCTCAAAAATTAAGGTTTATTAACAA 239
 550 GCCGAGATTAAGGAGGATTCACAGTTGGGCTCAAAAATTAAGGTTTATTAACAA 609
 240 GGGCAATTAAGTGGGCAATTCATTAAGCAAAAGCCCGATGATTTTCTGTAGTCA 299
 610 GGGCAATTAAGTGGGCAATTCATTAAGCAAAAGCCCGATGATTTTCTGTAGTCA 669
 300 GGTAAAGGCTGGGCACTTGGTGAATTAATTAATTAATTAATTAATTAATTAATTA 359
 670 GGTAAAGGCTGGGCACTTGGTGAATTAATTAATTAATTAATTAATTAATTAATTA 729
 360 GGTAAAGGCTGGGCACTTGGTGAATTAATTAATTAATTAATTAATTAATTAATTA 419
 730 GGTAAAGGCTGGGCACTTGGTGAATTAATTAATTAATTAATTAATTAATTAATTA 789
 420 TATAGAGTGTAAAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 479
 790 TATAGAGTGTAAAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 849
 480 TACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 539
 850 TACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 909
 540 TACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 599
 910 TACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 969
 600 TCTGAGGAGGAGTTTGGGAAATGATCAAGCAAAAGGCGCAAA----- 644
 970 TCTGAGGAGGAGTTTGGGAAATGATCAAGCAAAAGGCGCAAA----- 1029
 645 GTTGCAGGTCATATCATTAATTCGACGCTGCAATACACATCAATCAAGCTGACAGT 704
 1030 GTTGCAGGTCATATCATTAATTCGACGCTGCAATACACATCAATCAAGCTGACAGT 1089
 705 AATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 764
 1090 AATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1149
 765 AATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 824
 1150 AATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1209
 825 AATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 884
 1210 AATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1269
 885 TTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 941
 1270 TTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1329
 942 TCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1001
 1330 TCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1389

1002 GGGCTATTAATC-----AGAAATCAGAAATTCATCAGAAATTAATTAATTAATTA 1052
 1390 GGTAGCTTACTCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1449
 1053 GCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1112
 1450 GCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1509
 1113 GCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1172
 1510 GCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1569
 1173 CAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1232
 1570 CAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1629
 1233 GGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1292
 1630 GGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1689
 1293 AGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1352
 1690 AGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1749
 1353 ATTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1412
 1750 ATTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1809
 1413 GGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1472
 1810 GGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1869
 1473 AGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1532
 1870 AGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1929
 1533 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1592
 1930 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1989
 1593 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1652
 1990 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2049
 1653 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1709
 2050 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2109
 1710 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1769
 2110 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2169
 1770 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1829
 2170 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2229
 1830 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1889
 2230 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2289
 1890 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1949
 2290 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2343
 1950 TCGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2009
 2344 TCGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2403
 2010 TCGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2069
 2404 TCGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2463

QY 2070 TCTCAATTGAGGAAATTGGACAGTCAGCAATTAATGCAAAATGCAATTTGGTGTG 2129
 |||||
 Db 2464 TCTTCAATTGAGGAAATTGGACAGTCAGCAATTAATGCAAAATGCAATTTGGTGTG 2523
 |||||
 QY 2130 CCAATCAAGAAATACCAATTGACGCGTTCAAGATTGACAGGATTAACGACTTGTCA 2189
 |||||
 Db 2524 CCAATCAAGAAATACCAATTGACGCGTTCAAGATTGACAGGATTAACGACTTGTCA 2583
 |||||
 QY 2190 AAGTGAATTAAACGATACCAAAAGTTATTATTTATTCATACCAAAACCAATCAATGTC 2249
 |||||
 Db 2584 ACTGGAATTAAACGATACCAAAAGTTATTATTCATACCAAAACCAATCAATGTC 2643
 |||||
 QY 2250 TCTATTATTATTCGATATGCAACGCGCAATGTTAAAGTTAGCAAACTTAATGTC 2309
 |||||
 Db 2644 TCTATTATTATTCGATATGCAACGCGCAATGTTATTCATGTTTAGCAAACTTAATGTC 2703
 |||||
 QY 2310 AATGTCATTTAAACCAATCAAGCAATTTACATTTAGCAACCAATGCAACCAATGTC 2369
 |||||
 Db 2704 AATGTCATTTAAATCAATCAAGCAATTTACATTTAGCAACCAATGCAACCAATGTC 2763
 |||||
 QY 2370 AATATTCGACTTTCCGACATTTCACTGCAACGCGTGATATGCAAACTTTAGCGTAT 2429
 |||||
 Db 2764 AATATTCGACTTTCAATCAAGCAATTTCAAGCGTGATATGCAAACTTTAGCGTAT 2823
 |||||
 QY 2430 GTGCATTTAAGGATTCAGCTCAATTTCTTTAAAAACAGCCATTTTTCGACCAAT 2489
 |||||
 Db 2824 GTGCATTTAAGGATTCAGCTCAATTTCTTTAAAAACAGCCATTTTTCGACCAAT 2883
 |||||
 QY 2490 CAGGAGACAAAGCAACAGTGAAGTTGAAAATGCACTTGGCAATTCCTAGCGAT 2549
 |||||
 Db 2884 CAGGAGACAAAGCAACAGTGAAGTTGAAAATGCACTTGGCAATTCCTAGCGAT 2943
 |||||
 QY 2550 ACTATTCGAGATTTTAAGCTTAATTAACATGATGATCAGTTAAATCAGCTTAATTC 2609
 |||||
 Db 2944 GCGCATTTGAGATTTTAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTC 3003
 |||||
 QY 2610 GCTAGCTCAACCAATACGCGCAGT---CGCGCTATTAGAGCGGAAACAGCGCAAC 2666
 |||||
 Db 3004 GCTAGCTCAACCAATACGCGCAGT---CGCGCTATTAGAGCGGAAACAGCGCAAC 3063
 |||||
 QY 2667 TCGCAGACATGTTTCAACATTTGACATGACATGATGTAATGTAATGTAATGTAATG 2726
 |||||
 Db 3064 TCGCAGACATGTTTCAACATTTGACATGACATGATGTAATGTAATGTAATGTAATG 3123
 |||||
 QY 2727 TTTCAATTATCTTCAATCTTTATTTGCTAATTAAGGATTAATTTAAATTTATTC 2786
 |||||
 Db 3124 TTTCAATTATCTTCAATCTTTATTTGCTAATTAAGGATTAATTTAAATTTATTC 3183
 |||||
 QY 2787 GCTGAGGCGATTAATATTTATCTGTTGCAACAGCGCAACAGCCGAAACCTTTGAG 2846
 |||||
 Db 3184 GCTGAGGCGATTAATATTTATCTGTTGCAACAGCGCAACAGCCGAAACCTTTGAG 3243
 |||||
 QY 2847 CAATTAACTTTGTTGAAGCAAGATTAATCAACGTTATCAATGATTAATGTAATTTACT 2906
 |||||
 Db 3244 CAATTAACTTTGTTGAAGCAAGATTAATCAACGTTATCAATGATTAATGTAATTTACT 3303
 |||||
 QY 2907 TTAAGAAATGCAAGTTGATGCGAGCAATTAAGTTAATTTAGTAAGAAATGATGCG 2966
 |||||
 Db 3304 TTAAGAAATGCAAGTTGATGCGAGCAATTAAGTTAATTTAGTAAGAAATGATGCG 3363
 |||||
 QY 2967 GAATTCGCTTGCATACCAATTAAGCAAGCAAGAAATGCAATGATTTAGTAAGCA 3026
 |||||
 Db 3364 GAATTCGCTTGCATACCAATTAAGCAAGCAAGAAATGCAATGATTTAGTAAGCA 3423
 |||||
 QY 3027 GAGCAAGCAAGCAACATTTAGAGCAACAGTTGAACGACTGCTAAACCAACCA 3086
 |||||
 Db 3424 GAGCAAGCAAGCAACATTTAGAGCAACAGTTGAACGACTGCTAAACCAACCA 3483
 |||||
 QY 3087 GGTAGCAAGCAAGTTGCTCAAGAGCAAGG----- 3119
 |||||
 Db 3484 AGTATGCAAGAGTGGTCAAAAGCGGTTTCTGATACCTGCTGATCAAGC 3543
 |||||
 QY 3120 ----- 3119
 |||||

Db 3544 CAGTTAGAGTATTACAAAGCCGAACAGTTGACCGACTGCTGAAAAACAAAAAATTAAG 3603
 |||||
 QY 3120 -----NAGCAGCGTTTCTGATACCTGCTGATCAAAAGCTG 3158
 |||||
 Db 3604 GCAAAAAAGTCCGCTCAAAAAGAGCGGTGTTTCTGATACCTGCTGATCAAAAGCTG 3663
 |||||
 QY 3159 TTTAAACGATTAAGAAAGCAACAGCTGATGCTGAAACCAAAAAAGTAAGGCA 3218
 |||||
 Db 3664 TTAGAGTATTAAGAGCAAGCTGTAACCGCTGTAACCAAAAAATTAAGCA 3723
 |||||
 QY 3219 AAAACAAAAAGTCCGCTCAAAAAGAGCGTGTGTTTCTGATACCTGCTGATCAAAAG 3278
 |||||
 Db 3724 AAAAAGTCCGCTCAAAAAGAGCGCGGAGAGATTTTCTGATACCTGCTGATCAAAAG 3783
 |||||
 QY 3279 CTGTC---GCAATTGAGAGCGGCACTTGAGTTTATGATGCCCAACAGCAATGCAAAA 3335
 |||||
 Db 3784 CGGTTAAAGTATTAAGAGCAAACTTGAGGTTATTAATGCCCAACAGCAATGCAAAA 3843
 |||||
 QY 3336 GATGCTAGCTCAAGAAAGAGCGGCAAAACAGCAACCAAAAAAGCTGATCAGCGT 3395
 |||||
 Db 3844 GAACTCAAGATCA-----GGAACAGCAACCAAAAAAGCTGATCAGCGT 3894
 |||||
 QY 3396 TATTCAATGATGCGTATCAGATTTATCTGCAACAGTAATGATATGCTTTCTGTTCA 3455
 |||||
 Db 3895 TATTCAATGATGCGTATCAGATTTATCTGCAACAGTAATGATATGCTTTCTGTTCA 3954
 |||||
 QY 3456 GATGAATTGATGCTGTTTGTGATCAACCAACATGCGCGTGTGACAAATATGCA 3515
 |||||
 Db 3955 GATGAATTGATGCTGTTTGTGATCAACCAACATGCGCGTGTGACAAATATGCA 4014
 |||||
 QY 3516 CAGATTAAGAGCGCTGATGATCTGATGCGTGTGCTTATCAGACAGCAAAAGCAAC 3575
 |||||
 Db 4015 CAGATTAAGAGCGCTGATGATCTGATGCGTGTGCTTATCAGACAGCAAAAGCAAC 4071
 |||||
 QY 3576 TTAAGTCAATTTGAGGATGCAAAAAGCTTATGATGACAGTAATGAGGAGCTTTCTG 3635
 |||||
 Db 4072 TTAAGTCAATTTGAGGATGCAAAAAGCTTATGATGACAGTAATGAGGAGCTTTCTG 4131
 |||||
 QY 3636 CATAGCGCTGATGATTAATCTTGTGATCAAGTATTAATTAATTAATTAATTAATTA 3695
 |||||
 Db 4132 CATAGCGCTGATGATTAATCTTGTGATCAAGTATTAATTAATTAATTAATTAATTA 4191
 |||||
 QY 3696 ATGTGCGGTTTGGCCAAATGCAATGAGGAGCAATTAATTAATTTGATGATGAGGAG 3755
 |||||
 Db 4192 ATGTGCGGTTTGGCCAAATGCAATGAGGAGCAATTAATTAATTTGATGATGAGGAG 4251
 |||||
 QY 3756 GGAATCAGTGAATTAATGCTGAGAAACAGCAAAAGCTTATGATGAGGAGAT 3815
 |||||
 Db 4252 GGAATCAGTGAATTAATGCTGAGAAACAGCAAAAGCTTATGATGAGGAGAT 4311
 |||||
 QY 3816 AATTATGCGTGAATGCAATTAATGATTCGTTTGGGCAATGAGGATTAATGATGAT 3875
 |||||
 Db 4312 AATTATGCGTGAATGCAATTAATGATTCGTTTGGGCAATGAGGATTAATGATGAT 4371
 |||||
 QY 3876 TTTGAGTAAATGCTTATTTATGCAAGTAATTAATTAATTAATTAATTAATTAATTA 3935
 |||||
 Db 4372 TTTGAGTAAATGCTTATTTATGCAAGTAATTAATTAATTAATTAATTAATTAATTA 4431
 |||||
 QY 3936 AAAAGCTTACCTTGCATTTATGCTATTAATGCTGAGCAATGAGTGAATTAATTA 3995
 |||||
 Db 4432 AAAAGCTTACCTTGCATTTATGCTATTAATGCTGAGCAATGAGTGAATTAATTA 4491
 |||||
 QY 3996 ACTTCGAGATTAATTAACAGCTGTAAGCTTATTTCTGCAATTAATGATGATGAT 4055
 |||||
 Db 4492 ACTTCGAGATTAATTAACAGCTGTAAGCTTATTTCTGCAATTAATGATGATGAT 4551
 |||||
 QY 4056 AAGCTTAAGTAAACCAAGTAAATCTCAAGGTTGCAACCAATTTGAGCGTAT 4115
 |||||
 Db 4552 AAGCTTAAGTAAACCAAGTAAATCTCAAGGTTGCAACCAATTTGAGCGTAT 4611
 |||||
 QY 4116 TGGCAAAAAAGAGTGGGATTTAAAGCAAAATTTTACATTTCCGCTTTATC 4175
 |||||

Db 4612 TGGCAAAAGAGGAGATTAAGCGAAATTTTACATTTCCAACTTCCTGTTTAT 4671
 Qy 4176 TCAAAATCTCAAGTTCCAACTGGCAACAGCAAAATGCGCGTGAATTTGGGCTAT 4235
 Db 4672 TCAAAATCTCAAGTTCCAACTGGCAACAGCAAAATGCGCGTGAATTTGGGCTAT 4731
 Qy 4236 CTTGGTAAATCAATATTTTATGTTATGATTAACAAGTGGTGGATGAGA 4295
 Db 4732 CTTGGTAAATCAATATTTTATGTTATGATTAACAAGTGGTGGATGAGA 4788
 Qy 4296 TCCCACTTTTATTTCCAAATAT 4319
 Db 4789 --CCACCTTTTATTTCCAAATAT 4810

RESULT 9

US-10-080-505-16
 ; Sequence 16, Application US/10080505
 ; Patent No. 6676948
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/FBI/DCP/DHR
 ; CURRENT APPLICATION NUMBER: US/10/080,505
 ; PRIORITY FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 16
 ; LENGTH: 4828
 ; TYPE: DNA
 ; ORGANISM: Haemophilus influenzae
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (313)..(4548)
 ; OTHER INFORMATION:
 US-10-080-505-16

Query Match 68.8%; Score 2970; DB 4; Length 4828;
 Best local Similarity 82.3%; Pred. No. 0;
 Matches 3581; Conservatism 0; Mismatches 645; Indels 124; Gaps 10;

Qy 1 TCAATAGCTTTTAACTA-GTATTTTAAAGAAAATTTCTTAAATTAACAT 59
 Db 253 TCAATAGCTTTTAACTA-GTATTTTAAAGAAAATTTCTTAAATTAACAT 312
 Qy 60 ATGAAAAAACTATTTGCTTAAATTTTAAAGCTTGCATTTCAATAGGATAGTA 113
 Db 313 ATGAAAAAACTATTTGCTTAAATTTTAAAGCTTGCATTTCAATAGGATAGTA 372
 Qy 120 TCGCAAGCTGGGCTGTCACATTAATTTTGGATGATTAACAAATTAATGCTGATTT 179
 Db 373 TCGCAAGCTGGGCTGTCACATTAATTTTGGATGATTAACAAATTAATGCTGATTT 432
 Qy 180 GCCGAGATTAAGGAGGATTCAGCTGGGCTCAAAATTAATGATTTAAACAATA 239
 Db 433 GCCGAGATTAAGGAGGATTCAGCTGGGCTCAAAATTAATGATTTAAACAATA 492
 Qy 240 GGGCAATTAAGTGGCAATCATATGCAAAAGCCCGATGATTTTCTGATGCTCA 299
 Db 493 GGAATTTTAAGTGGCAATCATATGCAAAAGCCCGATGATTTTCTGATGCTCA 552
 Qy 300 CCTAAAGGCTGGCAAGCTTGGTGAATTAATATATTTAGAGCTGGCAATAGTA 359
 Db 553 CCTAAAGGCTGGCAAGCTTGGTGAATTAATATATTTAGAGCTGGCAATAGTA 612
 Qy 360 GGAATTAAGTGGCAATTTTGGTGAAGGAGGAAACAACCCCGATGATGCTGTTTCT 419
 Db 613 GGAATTAAGTGGCAATTTTGGTGAAGGAGGAAACAACCCCGATGATGCTGTTTCT 672

Qy 420 TATTAATGTTTAAAGCAATTAATCAAAAAAGATTTTAACTCTTATGAGACAT 479
 Db 673 TATTAATGTTTAAAGCAATTAATTAATAAGATTAAGCTCTTATGAGAAAGAC 732
 Qy 480 TACCATTAATCAAGATTAATTAATTTGTTACAGAGCGCTTCAATGATATGACTTCG 539
 Db 733 TACCATTAATCAAGATTAATTTTACAGAGCGCTTCAATGATATGACTTCG 792
 Qy 540 AATTAATGAGGATCTTTTCAAGATGAGCAAAATTTTCAAGAGCGCTTCAATGAGG 599
 Db 793 AATTAATGAGGATCTTTTCAAGATGAGCAAAATTTTCAAGAGCGCTTCAATGAGG 852
 Qy 600 TCGGAGCGGATTTTGGCAATGATCAAGCAAAAGCGCAACCAAGTTGCGGATAT 659
 Db 853 TCGGAGCGGATTTTGGCAATGATCAAGCAAAAGCGCAACCAAGTTGCGGATAT 912
 Qy 660 CATTA-CTTGAAGCTGGCAATACACAAATGAGGAGGATTAATGATTTGAT 719
 Db 913 CATTA-CTTGAAGGAGGATTAACACAAATGAGGAGGATTAATGATTTGAT 972
 Qy 720 TGGGAGCGGATTTTGGCAATGATCAAGCAAAAGCGCAACCAAGTTGCGGATAT 779
 Db 973 TGGGAGCGGATTTTGGCAATGATCAAGCAAAAGCGCAACCAAGTTGCGGATAT 1032
 Qy 780 GGGGAGGATTTTGGCAATGATCAAGCAAAAGCGCAACCAAGTTGCGGATAT 839
 Db 1033 GGGGAGGATTTTGGCAATGATCAAGCAAAAGCGCAACCAAGTTGCGGATAT 1092
 Qy 840 GGGATTTTACGGGAAGCAACCTTTTGAAGGCAAAAGATTTGATTTGATTTGCT 899
 Db 1093 GGGATTTTACGGGAAGCAACCTTTTGAAGGCAAAAGATTTGATTTGATTTGCT 1152
 Qy 900 AATCTTATTTGATTAATTTTGAAGGATTTTGAAGGATTTTGAAGGATTTTGAAG 959
 Db 1153 AATCTTATTTGATTAATTTTGAAGGATTTTGAAGGATTTTGAAGGATTTTGAAG 1212
 Qy 960 GGTATGAGGATTTTGAAGGATTTTGAAGGATTTTGAAGGATTTTGAAGGATTTTGAAG 1019
 Db 1213 GGTATGAGGATTTTGAAGGATTTTGAAGGATTTTGAAGGATTTTGAAGGATTTTGAAG 1272
 Qy 1020 TCGGAATTCATCAAGATTAATTAATTTTGAAGGATTTTGAAGGATTTTGAAGGATTT 1079
 Db 1273 TCGGAATTCATCAAGATTAATTAATTTTGAAGGATTTTGAAGGATTTTGAAGGATTT 1332
 Qy 1080 AAGGATTAAGTTCATCAAGATTTTGAAGGATTTTGAAGGATTTTGAAGGATTTTGAAG 1139
 Db 1333 AAGGATTAAGTTCATCAAGATTTTGAAGGATTTTGAAGGATTTTGAAGGATTTTGAAG 1386
 Qy 1140 AATGAGGAAGCTATTTTATGATCAAAAGCAAAAGATTTTATGCTGATGATGAC 1199
 Db 1387 AATGAGGAAGCTATTTTATGATCAAAAGCAAAAGATTTTATGCTGATGATGAC 1446
 Qy 1200 ATTACCAAGGAGGAGGATTTTATGATCAAAAGCAAAAGATTTTATGCTGATGATGAC 1259
 Db 1447 ATTACCAAGGAGGAGGATTTTATGATCAAAAGCAAAAGATTTTATGCTGATGATGAC 1506
 Qy 1240 AACCAAGCTTGGCAAGGATTTTATGATCAAAAGCAAAAGATTTTATGCTGATGATGAC 1319
 Db 1507 AACCAAGCTTGGCAAGGATTTTATGATCAAAAGCAAAAGATTTTATGCTGATGATGAC 1566
 Qy 1320 GTTAATGAGGATTTTATGATCAAAAGCAAAAGATTTTATGCTGATGATGAC 1379
 Db 1567 GTTAATGAGGATTTTATGATCAAAAGCAAAAGATTTTATGCTGATGATGAC 1626
 Qy 1380 GCGAAAGGAGGATTTTATGATCAAAAGCAAAAGATTTTATGCTGATGATGAC 1439
 Db 1627 GCGAAAGGAGGATTTTATGATCAAAAGCAAAAGATTTTATGCTGATGATGAC 1686
 Qy 1440 CAGGAGAGGATTTTATGATCAAAAGCAAAAGATTTTATGCTGATGATGAC 1499
 Db 1687 CAGGAGAGGATTTTATGATCAAAAGCAAAAGATTTTATGCTGATGATGAC 1746
 Qy 1500 GGGAGCTTTCAATTTAAAGATTAATTAATTTTATGATCAAAAGATTTTATGCTGATGAT 1559

[illegible]

Db	2809	AGTACGTATACGTTAAATTACGTTATATGACGTACCTCAATTAATATGCGCAAGTACCGC	2868
Qy	2637	CGTTCATTAGACGCGAAACAAGCTCAATATGGAGAAACAATCGTTTCAACAATGACA	2896
Db	2869	CGTTCATTAGACGCGAAACAAGCTCAATATGGAGAAACAATCGTTTCAACAATGACA	2928
Qy	2697	GTAAATGTGAATATGATGTGAGGCAAGACATTCGAATTACTTATCTTATTTGGCTAT	2756
Db	2929	GTAAATGTGAATATGATGTGAGGCAAGACATTCGAATTACTTATCTTATTTGGCTAT	2988
Qy	2757	AAAGCGATTAATTAATTAATATTCATGACGCTGAGGCGCATTCATATATCTGTCCG	2816
Db	2993	AAAGCGATTAATTAATTAATATTCATGACGCTGAGGCGCATTCATATATCTGTCCG	3048
Qy	2817	AACACAGGCAAAACCCGAAACCGCTGGAGCAATTACTTTGGTGAAGCAAAATAT	2876
Db	3049	GACACAGGCAAAACCCGAAACCGCTGGAGCAATTACTTTATTTGAAGCTTGGATAT	3108
Qy	2877	CACCGCTTACAGATAGCTCAAAATTACCTTAGAAATGACACGTTGATCAGGTGCA	2936
Db	3109	CACCGCTTACAGATAGCTCAAAATTACCTTAGAAATGACACGTTGATCAGGTGCA	3168
Qy	2937	TTACGTTTAAATTAATGTAAGAAATGATGGGAAATTCGCTTGCATTAACCATTAAG	2996
Db	3169	TGGCGTTTAAATTAATGTAAGAAATGATGGGAAATTCGCTTGCATTAACCATTAAG	3228
Qy	2997	CAGAAATTGACATATATTAGTAAAGCAGACAGACGACGACGACATTTAGAACCA	3056
Db	3229	CAGAAATTGCGCAATATATTAGTAAAGCAGACGACGACGACGACATTTAGAACCA	3288
Qy	3057	CAAGTTGAACCGACTCTTAAACCAACAAGGTGACCCAAAGTGGGTGCAGAAAGACA	3116
Db	3289	CAAGTTGAACCGACTCTTAAACCAACAAGGTGACCCAAAGTGGGTGCAGAAAGACA	3342
Qy	3117	GGCAGAGAGAGTTCCTGATATCCCTCCGATCAAAAGCTGTTAAAGCATTAAGACC	3176
Db	3343	---AAGGCGGCTTTCTGATATCCCGCTGATCAAAAGCATTTAAAGCATTAAGACC	3399
Qy	3177	AAACAGCTGAACCTGATGCTGTAACAACA---AAATTAAGGCATAAAACAA	3231
Db	3400	GAACTCAGACAGATTAATGCCCAACAAGATGACGACGATGACAAATGACAAAGTA	3459
Qy	3232	TGCGGTCAAAAAGACA---GTGTTTCTGATCCCGCTGATTAACAAGCCGTGCGCA	3287
Db	3460	ACTGCACTTAACCAAAAGACGACGACGATTAACCACTCAAGTTAAAGCAATTTAGTC	3519
Qy	3288	TTAGAGCGCGACTGATGATTTATGATCCCAAC-----	3321
Db	3520	TTGGCAACTGATGATGGTGGAAAAAGAAACCGCTCAGATGATTTGCTATGCAAAATTA	3579
Qy	3322	-----AGCAATCGAA	3332
Db	3580	GCTCAGTTGAATTTAACACACACTGAAAAAGCTTAGAGTGAGCTGGACGACGACAA	3619
Qy	3333	AAAGATCGTTCAGCTCAAGAAAGACGCGAAAAACAACGAAACAAATAGACTGATGAC	3392
Db	3640	AAAGAGCTTAAGCTCAAGAGACGCGAAAAACAACGAAACAAATAGACTGATGAC	3699
Qy	3393	CGTATTCAAATATGTGCTTATCAGAAATATCTGCAACAGTAAATAGATATGCTTCTGTT	3452
Db	3700	CGTATTCAAATATGTGCTTATCAGAAATATCTGCAACAGTAAATAGATATGCTTCTGTT	3759
Qy	3453	CAGAGTAATTAAGTGTGCTTTTGTGATCAGACCAACTGCGCGGTGGAACAATATC	3512
Db	3760	CAGAGTAATTAAGTGTGCTTTTGTGATCAGACCAACTGCGCGGTGGAACAATATC	3819
Qy	3513	GCAAGAGATTAAGACGCTATGATTTGATGCTTCGTTGCTTATCAGACGACGAAACG	3572
Db	3820	TGACAGATTAAGACGCTATGATTTGATGCTTCGTTGCTTATCAGACGACGAAACG	3876
Qy	3573	AACCTTAAGTCAATATGGGTGCAAAAGCCTTAGCAATGACAACTTGGGGCAGTTTC	3632
Db	3877	AACCTTAAGTCAATATGGGTGCAAAAGCCTTAGCAATGACAACTTGGGGCAGTTTC	3926

```

QY 3633 TCGCATAGCCGTCGATATATACCTTGTATGAACAGTTAAATACAGGACATTACG 3692
DB 3937 TCGCATAGCCGTCGATATATACCTTGTATGAACAGTTAAATACAGGACATTACG 3996
QY 3693 ATGATGTCGGGTTTTCGCCAATATCAATGAGGCGATTTTCAATTTGGTGAACGTCGA 3752
DB 3997 ATGATGTCGGGTTTTCGCCAATATCAATGAGGCGATTTTCAATTTGGTGAACGTCGA 4056
QY 3753 ACGGGAATCAGTCCGATTAATGCTGAGAAACAACCGAATAATTCCTGAAAACG 3812
DB 4057 ACGGGAATGTCGATTAATGCTGAGAAACAACCGAATAATTCCTGAAAACG 4116
QY 3813 ATAAATATGCGCGAATGCAAGTTATCACTTCGTTAAGGCAATGGGCAATTCAGCT 3872
DB 4117 ATAAATATGCGCGAATGCAAGTTATCACTTCGTTAAGGCAATGGGCAATTCAGCT 4176
QY 3873 TATTTGAGTATATGCTATTTTATTTGAACGTGAATATCAATCTGAGAGTGA 3932
DB 4177 TATTTGAGTATATGCTATTTTATTTGAACGTGAATATCAATCTGAGAGTGA 4236
QY 3933 GTGAAAACGCTTACCTTGCATTTATGCTATATGCTGCAATTGAGTATATCA 3992
DB 4237 GTGAAAACGCTTACCTTGCATTTATGCTATATGCTGCAATTGAGTATATCA 4296
QY 3993 TTTACTCGGACAGATATATCAAGGCTTACCTTATTTCTTGTCAATTATGAGTAT 4052
DB 4297 TTTACTCGGACAGATATATCAAGGCTTACCTTATTTCTTGTCAATTATGAGTAT 4356
QY 4053 TCAAAAGCTATACGTAACAACGATTAATCTCAAGGCTTGAACAACCACTTGGACGT 4112
DB 4357 TCAAAAGCTATACGTAACAACGATTAATCTCAAGGCTTGAACAACCACTTGGACGT 4416
QY 4113 TATTGGCAAAAAGATGAGATTAAGGCAAAATTTTACCTTCCAAATTTCCGCTTT 4172
DB 4417 TATTGGCAAAAAGATGAGATTAAGGCAAAATTTTACCTTCCAAATTTCCGCTTT 4476
QY 4173 ATCTCAAAATCTCAAGGCTTCAACCTGCGCAACAGCAAAATGAGGCGTGAATTTGG 4232
DB 4477 ATCTCAAAATCTCAAGGCTTCAACCTGCGCAACAGCAAAATGAGGCGTGAATTTGG 4536
QY 4233 TATCGTTGGTAAATCAACATTAATTTAT 4262
DB 4537 TATCGTTGGTAAATCAACATTAATTTAT 4566

```

RESULT 10
US-10-080-505-14

```

; Sequence 14, Application US/10080505
; Patent No. 6676948
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph V.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RT/DCF/DHR
; CURRENT APPLICATION NUMBER: US/10/080,505
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1.
; SEQ ID NO 14
; LENGTH: 4822
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (388)..(4563)
; OTHER INFORMATION:
; US-10-080-505-14

```

Query Match 64.1%; Score 2770.2; DB 4; Length 4822;

Best Local Similarity 79.6%; Pred. No. 0;
Matches 3464; Conservative 0; Mismatches 813; Indels 76; Gaps 13;

```

QY 1 TCAATATGCTTTTACCTA-GTATTTTATATGCAAAATTTCTTAATTAATTAACAT 59
DB 328 TCAATATGCTTTTACCTA-GTATTTTATATGCAAAATTTCTTAATTAATTAACAT 387
QY 60 ATGAAAAAAGCTGTTTGTCTTAATTTTAAACGCTGCAATTTCTATGAGGATATGA 119
DB 388 ATGAAAAAAGCTGTTTGTCTTAATTTTAAACGCTGCAATTTCTATGAGGATATGA 447
QY 120 TCGCAAGCTGGGCTGCTCACTTATTTTGGATGATTAACAAATTAATGAGTATTT 179
DB 448 TCGCAAGCTGGGCTGCTCACTTATTTTGGATGATTAACAAATTAATGAGTATTT 507
QY 180 GCGGAGATTAAGGAAATTCACAGTTGGGCTCAAAATTTTAAAGTTTAAACAAACA 239
DB 508 GCGGAGATTAAGGAAATTCACAGTTGGGCTCAAAATTTTAAAGTTTAAACAAACA 567
QY 240 GGGCAATTAATGAGCATCATCAATGACAAAGCCCGATGATTTTCTGTAGTGTCA 299
DB 568 GGGCAATTAATGAGCATCATCAATGAGCAAGTGTGCTTATGCTATTTATCTTCAATG 627
QY 300 CGTACGCGGTGAGCAAGCTTGTGTAATATATATGATATATGAGGCAATTAACGTA 359
DB 628 CGTACGCGGTGAGCAAGCTTGTGTAATATATGATATATGAGGCAATTAACGTA 687
QY 360 GATATACAGATGTTGATTTTGTGCAAGAGGAAACAACCCGATCAATGCTTTACT 419
DB 688 GATATACAGATGTTGATTTTGTGCAAGAGGAAACAACCCGATCAATGCTTTACT 747
QY 420 TATTAATGTTAAAGCAATTAACATAAAATTAATTTATCTTATGAGGACAT 479
DB 748 TATTAATGTTAAAGCAATTAATTAATTAAGGCGTATATGACATTAATTAAT- 801
QY 480 TACCATATCAACGATTAATTAATTTCTTACAGAGCGGCTTCAATGATATGACTTG 539
DB 802 TACCATATCAACGATTAATTAATTTCTTACAGAGCGGCTTCAATGATATGACTTG 861
QY 540 AATATTAATGCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 599
DB 862 AATATTAATGCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 921
QY 600 TCGGACGAGCTTTTGGCGAATGATCAAGCAAAAGGC---GACCAAGTCCGAGTGA 656
DB 922 AGTGGGCAATGATGAGTGAAGAGCAATTAATTAATTAATTAATTAATTAATTAAT 981
QY 657 TATCATTAATGCAATGATGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 716
DB 982 GGTTCATGTTTATAGGTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1041
QY 717 TATTGGAGAGGAAATTTGCTTAAGCGGAGAAATGATGCTTACCAATGAGGCTGA 776
DB 1042 GAATTAATTAATGAGGAGGAAATTTGCTTAAGCGGAGAAATGATGCTTACCAATGAG 1101
QY 777 AAGGGGAGAGGAGTGTCTCCGATGTTTATTAATTAATTAATTAATTAATTAATTAAT 836
DB 1102 TTGGGGAGAGGAGTGTCTCCGATGTTTATTAATTAATTAATTAATTAATTAATTAAT 1161
QY 837 AATGGGATTTTACGAGGAGGAAACCTTTTGAAGGCAAAAGAAATGAGTTTCAATTT 896
DB 1162 AATGGGATTTTACGAGGAGGAAATCCTTATGCTGCAAGGAAAGAGGATTAATTAATTA 1221
QY 897 GCGAATCTTATTTGATGAAATTTTCAAGGAGATTAATTAATTAATTAATTAATTAAT 956
DB 1222 GCGAATCTTATTTGATGAAATTTTCAAGGAGATTAATTAATTAATTAATTAATTAAT 1281
QY 957 GCTGTATATGAGTGTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1006
DB 1282 AATGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1341
QY 1007 -TATTAATGAGAAATTAAGGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1064

```

Db 1342 AGTACATTACTTAAAAAGCTATTCAGCCGTAACGAAATAGGGCTCTTTATGATTAATAGC 1401
 Qy 1065 TTACCTTGAAAGAGAAATPAA-----GTTCAATACTGATATAGGAGCCTAAT 1118
 Db 1402 CAATTCATGATGCTAGAGATTAATAATGCGATGATATCCCTTATTAAGGTCTAAT 1461
 Qy 1119 ATTATCTCCAGGTTTAAACATGAGAAAGCCTAATTTTATGATCAAAAACAGA 1178
 Db 1462 CCAATGCGCAGCATTAACATGAGAAAGATTAATCTTGGCAGTACAGAAACAGGA 1521
 Qy 1179 TCATTAATCTTGCACTGACATTAACAGAGGCGGCTGCTTATTTGAGGTAAT 1238
 Db 1522 ACTTAAACATTTGAAATTAATTAATCAAGGTGAGGTGATGATTTGAGGTAAT 1581
 Qy 1239 TTAAAGTATTCCTCAAT---TCTAACCAATCTGGGCAAGAGCTGGCATACATTAAGT 1295
 Db 1582 TTGTTGTAAAGCAATCAAAATTAATTAATTAATCTGGCAAGGTGAGGTGCTGTGGA 1641
 Qy 1296 GAAATATGACCGCTTACTTGGAAAGTAAATGCGTGGACATGATGACCTTTCTTAATTT 1355
 Db 1642 GAAGAAATGATCTGTTGAATGGCAGGTGATATCGAAGGCGATGCTTATCCAAATTT 1701
 Qy 1356 GGTAAAGGAACTGCACTTCAAGCCAAAGGGGAAATTAAGGTTGATGAGGTAGC 1415
 Db 1702 GGGCTGGAAACCTTACTGTTAATGTTAAAGGAAATTAAGGAAAGCTGAGTGTGGT 1761
 Qy 1416 GATGTAAGTCAATTTGGAGCAGGCGAGCATCAAGGCAACCAAGCCTTAAGT 1475
 Db 1762 AAGGTTGGTTGTGTGATCAACAGAGATGATGATGATGATCAAAAGCCTTTAA 1821
 Qy 1476 GAAATGGCTGGTTAGCGGCAAGGAGCTTTCAATTAACGATGATTAACATTTGAT 1535
 Db 1822 GAAGTTGGATGTAAGTGTAGAGTACCGTTCACTAAATGTCAGATCAAGTTGAT 1881
 Qy 1536 ACCGATTAATTTTATTTGCTTGGCTTTGCTGTGCTGCTTAATCTTAAGGCGATCTA 1595
 Db 1882 CTTAAACATTTTATTTGCTTGGCTTTGCTGTGCTTCAATCTTAATGAGGATCTTA 1941
 Qy 1596 ACCTTAAAGTATCCAAATTAAGGAGCGAGGCGCATGATGTAACCATTAATCAACT 1655
 Db 1942 AGCTTAAAGTATCCAAATTAAGGAGCGAGGCGCATGATGTAACCATTAATCAACT 2001
 Qy 1656 CAAGCCGCTAATGCTAATTAATGAGGAAAGCAAGCTTTCTAATTAAGGAAATTAAT 1715
 Db 2002 CAAGCCGCTAATTAATGATTAATGAGGAAAGCAAGCTTTCTAATTAATTAATTAAT 2049
 Qy 1716 ATTAATTAATGATTAATGAGGAAAGCAAGCTTTCTAATTAAGGAAATTAAT 1775
 Db 2050 AGCAAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2109
 Qy 1776 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1835
 Db 2110 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2169
 Qy 1836 TTGCTACTTCAAGTGTGATCAAAATTTAAAGGCGATTAATCCAAACAAAGGTTAACT 1895
 Db 2170 TTGCTACTTCAAGTGTGATCAAAATTTAAAGGCGATTAATCCAAACAAAGGTTAACT 2229
 Qy 1896 TTTTTCAGCGGTGAGCGGACGCGGCGCTCAATCAATTAATTAATTAATTAATTAAT 1955
 Db 2230 GTTTTGT 2289
 Qy 1956 ATGGAAGTATTAACCAAGGCGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2015
 Db 2290 ATGGAAGTATTAACCAAGGCGAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2249
 Qy 2016 AAAGCTGAAAACCTTCAAAATTTAAAGGCGAAATGTGTGTGTGTGTGTGTGTGTGTGT 2075
 Db 2350 AAAGCTGAAAACCTTCAAAATTTAAAGGCGAAATGTGTGTGTGTGTGTGTGTGTGTGT 2409
 Qy 2076 ATTAAGGAAATTTGAGCGTCAAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 2135
 Db 2410 ATTAAGGAAATTTGAGCGTCAAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 2469

Qy 2136 CAACAAATTAACATTTGACAGCGGTCAGATTTGACAGATTAACAGCTTGTCAAAAGTG 2195
 Db 2470 CAGCAAAATTAACATTTGACAGCGGTCAGATTTGACAGATTAACAGCTTGTCAAAAGTG 2529
 Qy 2196 GATTTAACCGTACAAAGTATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAAT 2255
 Db 2530 GATTTAACCGTACAAAGTATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAAT 2589
 Qy 2256 AATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2315
 Db 2590 AATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2649
 Qy 2216 ACTTAAACATTAACAGGCGCAATTTATTAAGCAACAAAGGCGCAATTAATTAATTAAT 2375
 Db 2650 ACTTAAACATTAACAGGCGCAATTTATTAAGCAACAAAGGCGCAATTAATTAATTAATTAAT 2709
 Qy 2376 CAGCTTTCCGACAAATTAACAGGCGGATTAATTAATTAATTAATTAATTAATTAATTAAT 2435
 Db 2710 AATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2769
 Qy 2436 TTAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2495
 Db 2770 TTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2829
 Qy 2496 GACAAAGGCAACAGTGAAGTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2555
 Db 2830 GAGGAAAGCAACAGTGAAGTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2889
 Qy 2556 TTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2615
 Db 2890 TTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2949
 Qy 2616 TGAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2669
 Db 2950 TGAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3009
 Qy 2670 GAGGAAAGCAACAGTGAAGTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2729
 Db 3010 GAGGAAAGCAACAGTGAAGTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3069
 Qy 2730 CAATTTACTTCAATTTTATTTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2789
 Db 3070 CAATTTACTTCAATTTTATTTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3129
 Qy 2790 GAGGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2849
 Db 3130 GAGGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3189
 Qy 2850 TTAATTTGCTAATTTTATTTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2909
 Db 3190 TTAATTTGCTAATTTTATTTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3249
 Qy 2910 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2969
 Db 3250 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3309
 Qy 2970 TTCCGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3029
 Db 3310 TTCCGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3369
 Qy 3030 CAAGGAAAGCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3089
 Db 3370 CAAGGAAAGCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3429
 Qy 3090 GAGGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3149
 Db 3430 GAGGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3480
 Qy 3150 CAATTTACTTCAATTTTATTTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3209
 Db 3481 CAATTTACTTCAATTTTATTTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3537

QY	3210	AGTAAAGCAAAAAA	CAAAAAA	ATGGCGTCAAAA	AGAGAGGTCTTGATGCCCTGCTT	3259
Dz	3538	AGTAAAGCAAAAAA	GTGGCGTCAAAA	AGACTCGAGAGAGTCTTGATGCCCTGCTT	3597	
QY	3270	GATCAAAAGCTGTTT	CGCATTTAGAACCCGCACTTGAGATTATGATGCCCA	CACATCG	33229	
Dz	3598	GATCAAA-----	TATTACAGCCGCACTTGAGATTATGATGCCCA	CACAGT	36488	
QY	333C	GAATAAGATCGTCTAGCTCA--	GAAGAGCGAAAAA	CAAGCAAAAA	GACTTG	33868
Dz	3649	AAAAAAGAACTCAACT	CAAGAGAGAAAAA	CAAGCAAAAA	GAAATTG	37088
QY	3387	ATCAAGCTTATCAAAATG	CGTTATCGATTATCTGCAACAGTAAATATGATGCTT	34486		
Dz	3709	ATCAAGCTTATCAAAATG	CGTTATCGATTATCTGCAACAGTAAATATGATGCTT	37868		
QY	3447	TCTGTTCAAGATGATTAAT	GAATGCTCTTTTGTATGATCAAGCAATCTGCGTGGACA	35068		
Dz	3769	TGCGTTCAAGATGATGATGATGCTCTTTTGTATGATCAAGCAATCTGCGTGGACA	38288			
QY	3507	AATATCGCAACGATATAAAG	CGATATGATTCATGATGCTTGGCTTATCAAGCAAG	35668		
Dz	3829	AATATCGCAACGATATAAAG	CGCTATGATTCATGATGCTTGGCTTAT--CAAG	38858		
QY	3567	AAAAAGAACTTACGTCAAA	TTGAGGTGCAAAAAGCTTGATATGACAAATGAGGACA	36268		
Dz	3886	AAAAAGAACTTACGTCAAA	TTGAGGTGCAAAAAGCTTGATATGACAAATGAGGACA	39458		
QY	3627	GTGTTGCGCAACCGCTCAAT	TAATACCTTATATGACAGGTATATAACAGGACA	36868		
Dz	3946	GTGTTGCGCAACCGCTCAAT	TAATACCTTATATGACAGGTATATAACAGGACA	40058		
QY	3687	TTAACAGATATGTGGAGTTT	TGCGCAATCATAGGCGCATTTACAAATTGCTATTAAC	37468		
Dz	4006	TTAACAGATATGTGGAGTTT	TGCGCAATCATAGGCGCATTTACAAATTGCTATTAAC	40558		
QY	3747	GTGGAACCGGAAATCAT	GTGCAATTAATGCTGAAAGCAAAAGCCGAAATTCATCGA	38068		
Dz	4066	GTGGAACCGGAAATCAT	GTGCAATTAATGCTGAAAGCAAAAGCCGAAATTCATCGA	41258		
QY	3807	AAAGCAATTAATTTAGCGTGA	ATCCAGTTACGTTACGTTAGGCGAATAGGCAAT	38668		
Dz	4126	AAAGCAATTAATTTAGCGTGA	ATCCAGTTACGTTACGTTAGGCGAATAGGCAAT	41858		
QY	3867	CAGCTTATTTGAGATTAAT	CGTATTTATGAACTGAAATTAATCATCTAGAGA	39268		
Dz	4186	CAGCTTATTTGAGATTAAT	CGTATTTATGAACTGAAATTAATCATCTAGAGA	42458		
QY	3927	GTGAGAGTAAAGCGCTAG	CTGCAATTAATGCTATATGCTAGGCAATGAGTGAAT	39868		
Dz	4246	GTGAGAGTAAAGCGCTAG	CTGCAATTAATGCTATATGCTAGGCAATGAGTGAAT	43058		
QY	3987	TATACATTTACTCGACAG	TATATATCAAGGTAAAGCTTATTTCTGCTCAATTAATG	40468		
Dz	4306	TATACATTTACTCGACAG	TATATATCAAGGTAAAGCTTATTTCTGCTCAATTAATG	43658		
QY	4047	GATGTTCAAAAGCTAAAT	CGTAAACACGCTTAATCTCAGGTGTGACAAACATTT	41068		
Dz	4366	GATGTTCAAAAGCTAAAT	CGTAAACACGCTTAATCTCAGGTGTGACAAACATTT	44258		
QY	4107	GAAAGTATTTGGCAAAA	AGAGTGGATTAAGAGCAAAATTTTACATTCGAATTTCC	41668		
Dz	4426	GAAAGTATTTGGCAAAA	AGAGTGGATTAAGAGCAAAATTTTACATTCGAATTTCC	44858		
QY	4167	GCTTTATCTCAAAATCT	CAAGGTCAAACTGGCAAAAGCAAAATGAGGCTGAGAA	42268		
Dz	4486	GCTTTATCTCAAAATCT	CAAGGTCAAACTGGCAAAAGCAAAATGAGGCTGAGAA	45458		
QY	4227	TTGGAGTATCGTGTATTA	AAATCAAAATTTTATGCTTTATGTAATAA	CAAGGTGAGT	42868	
Dz	4546	TTGGAGTATCGTGTATTA	AAATCAAAATTTTATGCTTTATGTAATAA	CAAGGTGAGT	46058	
QY	4287	CAAGTCAAGTCCACCTTT	TATTCATCAATTAAT	4319		

[illegible]


```

Db 3104 CATTTCTTCATCTTTATTTGGCTATTAAGGATTAATTAATATCAATCAAGCT 3163
Qy 2796 GAGGGCGATTACATTAATATCTGTCGCAACAGCGAAGAACCCGAAACCTTGAGAA 2849
Db 3164 GAGGGCGATTACATTAATCTGTCGCAACAGCGAAGAACCCGAAACCTTGAGAA 3223
Qy 2850 TTAATCTTGGTTAAAGCAAGATATCAACCGTTATCAGATTAAGCTCAATTTCTTA 2909
Db 3224 TTAATCTTGGTTAAAGCAAGATATCAACCGTTATCAGATTAAGCTCAATTTCTTA 3283
Qy 2910 GAAATGAACAGCTGATGACAGCTGATGACAGCTGATGACAGCTGATGACAGCT 2969
Db 3284 GAAATGAACAGCTGATGACAGCTGATGACAGCTGATGACAGCTGATGACAGCT 3343
Qy 2970 TTCCGCTTGCAATACCCATTAAGAGCAGAAATGCAATGATTTGTAAGACAGAG 3029
Db 3344 TTCCGCTTGCAATACCCATTAAGAGCAGAAATGCAATGATTTGTAAGACAGAG 3403
Qy 3030 CAGCAGACGACATTAAGAGCAGAAATGCAATGATTTGTAAGACAGAG 3089
Db 3404 CAGCAGACGACATTAAGAGCAGAAATGCAATGATTTGTAAGACAGAG 3463
Qy 3090 GAGCCAAAGTGGCTCAAGAAAGAGCAGAGCAGAGCTTCTGATACCTGCTGAT 3149
Db 3464 GAGCCAAAGTGGCTCAAGAAAGAGCAGAGCAGAGCTTCTGATACCTGCTGAT 3514
Qy 3150 CAAAGCCTGTTAAACCATTAAGAGCAGAAATGCAATGATTTGTAAGACAGAG 3209
Db 3515 CAAAGCCTGTTAAACCATTAAGAGCAGAAATGCAATGATTTGTAAGACAGAG 3571
Qy 3210 AGTAAGCGCAAAACCAAAAGAGCGCTCAAAAGAGCAGAGCTTCTGATACCTGCT 3269
Db 3572 AGTAAGCGCAAAACCAAAAGAGCGCTCAAAAGAGCAGAGCTTCTGATACCTGCT 3631
Qy 3270 GATCAAAAGCTGTTGCAATTAAGAGCAGAAATGCAATGATTTGTAAGACAGAG 3329
Db 3632 GATCAAAAT-----ATTACAGCGCAGCTGAGGTATGATGACCAAGCAGAGAGT 3682
Qy 3330 GAAAGAGTGTCTAGCTCA-----GAGAGCGGAGCAAAACCAAGCAGAAAGAGCT 3386
Db 3683 GAAAGAGTGTCTAGCTCA-----GAGAGCGGAGCAAAACCAAGCAGAAAGAGCT 3742
Qy 3387 ATCAGCGCTTATCAATAGTGCCTTATCAGATTAATGCAATGATTTGTAAGACAGAG 3446
Db 3743 ATCAGCGCTTATCAATAGTGCCTTATCAGATTAATGCAATGATTTGTAAGACAGAG 3802
Qy 3447 TCTGTTCAAGATTAATGATGCTTTTGTGATGACAGAGCAATGCTGCTGAGCA 3506
Db 3803 TCTGTTCAAGATTAATGATGCTTTTGTGATGACAGAGCAATGCTGCTGAGCA 3862
Qy 3507 AATATCGAGAGATTAAGAGCAGTATGATGATGATGATGATGATGATGATGATGAT 3566
Db 3863 AATATCGAGAGATTAAGAGCAGTATGATGATGATGATGATGATGATGATGATGAT 3919
Qy 3567 AAAAGCAATTAAGAGTGGAGTGAAGAAAGCTTATGCTAAATGAGCAATGAGAGCA 3626
Db 3920 AAAAGCAATTAAGAGTGGAGTGAAGAAAGCTTATGATTAATGAGCAATGAGAGCA 3979
Qy 3627 GTTTTCTGCAATGAGCTTCAATTAATGCTTATGATGAGCAATGAGTAAATGAGCA 3686
Db 3980 GTTTTCTGCAATGAGCTTCAATTAATGCTTATGATGAGCAATGAGTAAATGAGCA 4039
Qy 3687 TTAAGCAATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3746
Db 4040 TTAAGCAATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4099
Qy 3747 GTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3806
Db 4100 GTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4159
Qy 3807 AAAGGATTAATTAATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 3866

```

```

Db 4160 AAAGGATTAATTAATGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4219
Qy 3867 CAGCTTATTTGAGAGTAAATGCTATTTATTAAGTGAATTAATCAATCGAGAA 3926
Db 4220 CAGCTTATTTGAGAGTAAATGCTATTTATTAAGTGAATTAATCAATCGAGAA 4279
Qy 3927 GTGAGAGTGAAGAGCTTACCTTATTAATGCTATTAATGCTATTAATGCTATTAAT 3986
Db 4280 GTGAGAGTGAAGAGCTTACCTTATTAATGCTATTAATGCTATTAATGCTATTAAT 4339
Qy 3987 TATACATTTACTCGGAGATTAATCAAGCTTAAGCTTATTTCTGTAATTAATG 4046
Db 4340 TATACATTTACTCGGAGATTAATCAAGCTTAAGCTTATTTCTGTAATTAATG 4399
Qy 4047 GATGTTCAAAAGCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4106
Db 4400 GATGTTCAAAAGCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4459
Qy 4107 GAGAGTATTTGCAAAAGAGTGGATTAAGAGCAGAAATTTACATTTCAAAATTTCC 4166
Db 4460 GAGAGTATTTGCAAAAGAGTGGATTAAGAGCAGAAATTTACATTTCAAAATTTCC 4519
Qy 4167 GCTTTATCTCAAAATCTCAAGTTCAACTGCGCAAGAGCAAAATGCGCGTGA 4226
Db 4520 GCTTTATCTCAAAATCTCAAGTTCAACTGCGCAAGAGCAAAATGCGCGTGA 4579
Qy 4227 TTGGGCTATCTGTTGTTAAATCAAC 4252
Db 4580 TTGGGCTATCTGTTGTTAAATCAAC 4605

```

```

RESULT 12
US-08-210-535-5
; Sequence 5, Application US/08210535
; Patent No. 5965424
; GENERAL INFORMATION:
; APPLICANT: Ambrosius, Dorothea
; APPLICANT: Dony, Carola
; APPLICANT: Rudolph, Rainer
; TITLE OF INVENTION: Recombinant TGA Protease
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felle & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release.#1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/210,535
; FILING DATE:
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 07/820,701
; FILING DATE: 10-Jan-1992
; APPLICATION NUMBER: DE 41 00 704.2
; FILING DATE: 11-Jan-1991
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: DE 41 40 699.0
; FILING DATE: 10-Dec-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5965424man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:

```

SEQUENCE CHARACTERISTICS:
 LENGTH: 4899 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA
 US-08-210-535-5

Query Match 3.8%; Score 164.8; DB 2; Length 4899;
 Best Local Similarity 58.0%; Pred. No. 3,6e-32;
 Matches 311; Conservative 0; Mismatches 222; Indels 3; Gaps 1;

QY 1152 CTATATTTATGATCAAAAACAGATCATTAATCTTCGATGATGATCACTTAACCAAGG 1211
 DB 1202 CAATATGTTACCTTTGAGACACAGCACTTGTATTTGATCAAAAATCAACCAAGG 1261
 QY 1212 GCGGGTGTCTTTATTTTGAAGGTAATTTACAGT--ATCTCAAAATCTTAACCAACT 1268
 DB 1262 GCGGGGCGGCTGTTTTCAAGGCGATTCACAGTCAAAAGCGCAATATGATCACTACT 1321
 QY 1269 TGGCAGAGAGCTGGCATCATGTAAGTAAATAGACACCTTACTTTGAAATGAAATGGC 1328
 DB 1322 TGGTTAGGTGCGGGATTTGATGTCGCAAGCAAAAAGTCTTTGGCAAGTCAAAAT 1381
 QY 1329 GTGGAACATGATGCACTTCTTAATTTGTAAGGAACATGCAAGCTTCAAGCAAGG 1388
 DB 1382 CCGAATGCGACAGATTCGCAAAATTCGCAAAAGCACTTTGAAATTAAGCGCAAGC 1441
 QY 1389 GAAATTAAGTTGATGATGAGGCGATGCTAAAGTCAATTTGAGACAGAGCGAGAC 1448
 DB 1442 GTTAAACCAAGGCAATTAAGGTGCGCAAGCTACGCTTAATCTGATCAAAAACCGAT 1501
 QY 1449 GATTAAGCGCAAAACCAAGCTTTATGTAATTTGCTGTGAGCGCGAGGAGCTGTT 1508
 DB 1502 GCGCAAAAAGTTCAGGCTTCTCCCAAGTGGCAATGTGCGAGCGATCAATG 1561
 QY 1509 CAATTAACGATGATTAACCAATTTGATTAACCAATTTATTTGCGCTTTCGATGAT 1568
 DB 1562 GTATTAATTAATGTTCAATTCAGATTAATCCCGATTAACCTTAATTTGCGTTCG 1621
 QY 1569 CGCTTAATCTTAACGCGCATTCATTAACCTTTAAAGTTCMAAATTCGAGCGAGG 1628
 DB 1622 CGTTGATGCGCAATGCGCAATGATGATCTTTGAAACATCCGCAAGCTGGAAGA 1681
 QY 1629 GCATGATTTGTAACCAATTAACCACTCAAGCGCTAATGCTATTTCTCGGGA 1684
 DB 1682 GCGGCAATTTGTAACCAACCAACAGCAAGCGCTCCCAATCACTTAAGGTTA 1737

RESULT 13
 5268270-1
 Patent No. 5268270
 APPLICANT: Meyer, Thomas F.; Halter, Roman; Pohner, Johannes
 TITLE OF INVENTION: PROCESS FOR PRODUCING PROTEINS USING GRAM
 NEGATIVE HOST CELLS
 NUMBER OF SEQUENCES: 6
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/171,872
 FILING DATE: 01-JUL-1987
 SEQ ID NO: 1
 LENGTH: 4899
 5268270-1

Query Match 3.8%; Score 164.8; DB 6; Length 4899;
 Best Local Similarity 58.0%; Pred. No. 3,6e-32;
 Matches 311; Conservative 0; Mismatches 222; Indels 3; Gaps 1;

QY 1152 CTATATTTATGATCAAAAACAGATCATTAATCTTCGATGATGATCACTTAACCAAGG 1211
 DB 1202 CAATATGTTACCTTTGAGACACAGCACTTGTATTTGATCAAAAATCAACCAAGG 1261
 QY 1212 GCGGGTGTCTTTATTTTGAAGGTAATTTACAGT--ATCTCAAAATCTTAACCAACT 1268

DB 1262 GCGGCGCGCTGTTTTCMAAGGCGATTAACAGTCAAGTCAAAAGCGCAAAATATGATCACT 1321
 QY 1269 TGGCAAGAGAGCTGGCATTAATGATGTAAGTAAGTAAGCAAGCGTACTTGAAGTAATGGC 1328
 DB 1322 TGGTTAGGTGCGGGATTTGATGTCGCAAGCAAAAAGTCTTTGCAAGTCAAAAT 1381
 QY 1329 GTGGAACATGATGCACTTCTTAATTTGTAAGGAACATGCAAGCTTCAAGCAAGG 1388
 DB 1382 CCGAATGCGACAGATTTGCAAAAATTCGCAAAAGCACTTTGAAATTAAGCGCAAGC 1441
 QY 1389 GAAATTAAGTTGATGATGAGGCGATGCTAAAGTCAATTTGAGACAGAGCGAGAC 1448
 DB 1442 GTTAAACCAAGGCAATTAAGGTGCGCAAGCTACGCTTAATCTGATCAAAAACCGAT 1501
 QY 1449 GATTAAGCGCAAAACCAAGCTTTATGTAATTTGCTGTGAGCGCGAGGAGCTGTT 1508
 DB 1502 GCGCAAAAAGTTCAGGCTTCTCCCAAGTGGCAATTTGTAAGCGCGATCAATG 1561
 QY 1509 CAATTAACGATGATTAACCAATTTGATTAACCAATTTATTTGCGCTTTCGATGAT 1568
 DB 1562 GTATTAATTAATGTTCAATTCAGATTAATCCCGATTAACCTTAATTTGCGTTCG 1621
 QY 1569 CGCTTAATCTTAACGCGCATTCATTAACCTTTAAAGTTCMAAATTCGAGCGAGG 1628
 DB 1622 CGTTGATGCGCAATGCGCAATGATGATCTTTGAAACATCCGCAAGCTGGAAGA 1681
 QY 1629 GCATGATTTGTAACCAATTAACCACTCAAGCGCTAATGCTATTTACTGGA 1684
 DB 1682 GCGGCAATTTGTAACCAACCAACAGCAAGCGCTCCCAATCACTTAAGGTTA 1737

RESULT 14
 US-08-232-463-14/C
 Sequence 14, Application US/08232463
 Patent No. 5670367
 GENERAL INFORMATION:
 APPLICANT: DORNER, F.
 APPLICANT: SCHEIFLINGER, F.
 APPLICANT: FALKNER, F. G.
 TITLE OF INVENTION: RECOMBINANT FOMULOX VIRUS
 NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 1800 Diagonal Road, Suite 500
 CITY: Alexandria
 STATE: VA
 COUNTRY: USA
 ZIP: 22313-0299
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/232,463
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/935,313
 FILING DATE:
 APPLICATION NUMBER: EP 91 114 300.6
 FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 INVTJ
 TELEPHONE: (703) 836-9300
 TELEFAX: (703) 683-4109
 TELEX: 899143
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7216 base pairs

```

;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZpc-flis
; US-08-232-463-14

```

```

Query Match
Best Local Similarity 6.6%; Pred. No. 8.8e-06;
Matches 28; Conservative 227; Mismatches 170; Indels 0; Gaps 0;

```

```

QY 2851 TAACTTGGTTAAAGCAAGATATCAACCGTTATCGATTAAGCTCAATTACTTTAG 2910
DB 1474 TATCTATGCAAGTAGTAAAGAGATAGAAATTTGTACRRRRRRRRRRRRRRRRR 1415
QY 2911 AAAATGACCGTTGATGAGGTGATTAAGTTAAATTTGTGAAAGAAATGATGCCAT 2970
DB 1414 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1355
QY 2971 TCCGCTTGATACCAATTAAGAGAGAGATTCACATATGATTTAGTAGAGAGAGC 3030
DB 1354 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1295
QY 3031 AAGCAGACGACATTAGAACCCAAACAGTTGAACCGACTGTAACCAACAAACAGGTG 3090
DB 1294 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1235
QY 3091 ACCCAAAAGTGGCGTCAAGAAAGACAGCAGCGCTTCCGTGATACCTGCTGATC 3150
DB 1234 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1175
QY 3151 AAAGCTGTTAAAGCATTAGAGCCAAACAGCTGAAGTGAAGTGAACACAAACAAA 3210
DB 1174 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1115
QY 3211 GTAAGGCAAAAACAAAAAAGTGGCGTCAAAAAGACAGGTGTTTCTGATCCCTGCTTG 3270
DB 1114 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1055
QY 3271 ATCAA 3275
DB 1054 CTCGA 1050

```

RESULT 15

```

US-09-621-976-2813/C
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
; US-09-621-976-2813

```

```

Query Match
Best Local Similarity 1.1%; Score 48.6; DB 4; Length 832;
Matches 45; Conservative 175; Mismatches 169; Indels 0; Gaps 0;

```

```

QY 1510 AATTAAAGATGATAACCAATTGATACCATTAATTTATTTGGCTTCGAGTGGTC 1569
|||||

```

```

DB 390 AATTGACATTAATTAATTAATTTGTTGTTWTKWTTWTTWTTWTKKARWY 331
QY 1570 GCTTAGATCTTAACGGGCAATCATTAACCTTTAAAGTATCCAAATACGAGAGGGG 1629
DB 330 WWSSTYACASRYRYTWGWTWTRWSTWTCWCKCYRRCATYTAARWMS 271
QY 1630 CAATGATTTGTAACATTAATTAACACTCAAGCCGTAATGTCATTAATCTGGAGAGAA 1689
DB 270 YAMGWKMSRBMMSMCTRMWYKKSSTYWTMCTCATWCTWYKWKWMSKTSRGG 211
QY 1690 GCATGTTCTACTAATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1749
DB 210 YMTSYSTRSYSMATWMSWWTYCWMMGRWSTWYMAWGWKWRYYATWRAAMWMAW 151
QY 1750 ACAAGGTTGTTGGGAAACAGATAAAATAACCAATGAGGATTAACCTTAATTT 1809
DB 150 TMMWYWMWMSRGAATYRTRTWGTYRWKSYRTRCANAAYAKTKGSYTWK 91
QY 1810 ATAAACCAACAGAGAGATCGTACTTGTCTACTTTGAGTGTACAAATTTAAAGGGG 1869
DB 90 MKWKRGMWMAAAYKTMWRACTKTRWMAWMAWMAWMAWMAWMAWMAWMAWMAW 31
QY 1870 ATATTAACCAACAAAGTAAACTATTT 1898
DB 30 MKWKRMSWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAW 2

```

Search completed: June 24, 2004, 19:44:58
 Job time : 308 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2004, 17:29:18 ; Search time 28 Seconds
(without alignments)
4788.964 Million cell updates/sec

Title: US-10-645-655-2

Perfect score: 7273
Sequence: 1 MKKTFRLNPLTACTISIGIV.....SQGSOLGKQONVGLGYRW 1394

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 segs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_78:.*
1: p1r1:.*
2: p1r2:.*
3: p1r3:.*
4: p1r4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7273	100.0	1394	2 S60762	IGA-specific serin
2	4178	57.4	1449	2 B81963	IGA-specific serin
3	4085	56.2	1457	2 DB1019	adhesion and penet
4	2634.5	36.2	709	2 C64057	IGA-specific metal
5	1676	23.0	1431	2 A81018	serine-type peptid
6	1673.5	23.0	1541	2 A37023	IGA-specific metal
7	1662.5	22.9	1545	2 B41859	IGA-specific metal
8	1654.5	22.7	1561	2 S61314	IGA-specific metal
9	1617	22.2	1532	2 A26039	IGA-specific metal
10	1609	22.1	1634	2 H64106	IGA-specific metal
11	1605	22.1	1702	2 A41859	IGA-specific metal
12	1555.5	21.4	1849	2 C41859	IGA-specific metal
13	1544.5	21.2	1815	2 C81169	IGA-specific metal
14	1536.5	21.1	1773	2 A81937	IGA-specific metal
15	1113.5	15.3	1377	2 I54632	tsh protein - 3sch
16	935	12.9	1366	2 S57664	IGA-specific metal
17	847	11.6	508	2 S61319	IGA-specific metal
18	845.5	11.6	497	2 S61328	IGA-specific metal
19	836	11.5	508	2 S61329	IGA-specific metal
20	835	11.5	508	2 S61330	IGA-specific metal
21	824	11.3	486	2 S61337	IGA-specific metal
22	824	11.3	486	2 S61335	IGA-specific metal
23	823.5	11.3	507	2 S61331	IGA-specific metal
24	821	11.3	508	2 S61330	IGA-specific metal
25	819	11.3	496	2 S61332	IGA-specific metal
26	811	11.2	496	2 S61325	IGA-specific metal
27	806	11.1	482	2 S61332	IGA-specific metal
28	707	9.7	471	2 S61332	IGA-specific metal
29	678	9.3	471	2 S61333	IGA-specific metal

30	604	8.3	1300	2 T00317	probable serine pr
31	368.5	5.7	384	2 B41500	IGA-specific metal
32	344.5	4.7	1569	2 A65044	hypothetical prote
33	341.5	4.7	1528	2 D85912	hypothetical prote
34	341.5	4.7	1571	2 C91068	hypothetical prote
35	326.5	4.5	1430	2 AF0355	probable autotrans
36	313	4.3	949	2 T08658	hypothetical prote
37	292.5	4.0	375	2 AD0123	probable autotrans
38	288	4.0	836	2 D64988	yej0 protein - Bac
39	283.5	3.9	1536	2 A43855	high-molecular-we
40	281	3.9	836	2 A98014	hypothetical prote
41	281	3.9	836	2 C85858	hypothetical prote
42	280	3.8	1477	2 B43855	high-molecular-we
43	268.5	3.7	1327	2 B90674	Aida-I adhesin-lik
44	268.5	3.7	1349	2 E85524	probable beta-bar
45	258.5	3.6	1645	2 JN0896	crystalline surfac

ALIGNMENTS

RESULT 1

S60762 IGA-specific serine endopeptidase (EC 3.4.21.72) precursor - Haemophilus influenzae (str

N)/Alternate names: adhesion and penetration protein hap

C/Species: Haemophilus influenzae

C/Date: 27-Apr-1996 #sequence - revision 13-Mar-1997 #text change 08-Dec-2000

C/Accession: S60762; S78001

R/St Gene III, J.W. de la Morena, M.L.; Falkow, S.

Mol. Microbiol. 14, 217-233, 1994

A/Title: A Haemophilus influenzae IGA protease-like protein promotes intimate interacti

A/Reference number: S60762; MIMD:95131744; PMID:7830568

A/Accession: S60762

A/Status: nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: 1-1394 <STC>

A/Cross-references: EMBL:U11024

A/Experimental source: strain N187

A/Accession: S78001

A/Molecule type: protein

A/Residues: 27-33 <ST2>

C/Genetics:

A/Genes: hap

C/Superfamily: IGA-specific metalloendopeptidase

C/Keywords: hydrolase; serine proteinase

F/-26/Domain: signal sequence #status predicted <SIG>

F/27-1394/Product: IGA-specific serine endopeptidase #status experimental <MAT>

F/243/Active site: Ser #status predicted

Query Match 100.0%; Score 7273; DB 2; Length 1394;

Best local similarity 100.0%; Pred. No. 0;

Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKKTFRLNPLTACTISIGIVSQMAAGHYFGIDVYQYRPAENKKEFYVGNQIKYNNQ	60
DB	1	MKKTFRLNPLTACTISIGIVSQMAAGHYFGIDVYQYRPAENKKEFYVGNQIKYNNQ	60
QY	61	GOLVGSMTKAPMDPFSVSRNCAVALYENQYISVAHWGTYDVPFAGENNPDHRT	120
DB	61	GOLVGSMTKAPMDPFSVSRNCAVALYENQYISVAHWGTYDVPFAGENNPDHRT	120
QY	121	YKIVKNNYKKNLHPFEDYRERLEKVTLEAPDIMTSNNGSTYSDRYKRYERVRI	180
DB	121	YKIVKNNYKKNLHPFEDYRERLEKVTLEAPDIMTSNNGSTYSDRYKRYERVRI	180
QY	181	SGRQFNRNDQKQOVAGAYHTAGTNHNRGANGSYLGGDYRKAGEYGPPIAGSK	240
DB	181	SGRQFNRNDQKQOVAGAYHTAGTNHNRGANGSYLGGDYRKAGEYGPPIAGSK	240
QY	241	GDSGSMFYDAEKQKWLINGIIRREGNPFBEKENGFPOLVRKSYFDEIRDLHTSLYTR	300
DB	241	GDSGSMFYDAEKQKWLINGIIRREGNPFBEKENGFPOLVRKSYFDEIRDLHTSLYTR	300

Applied

QY 301 GNGVYTTISGNDGQSITQKSGIDPSEIKITLAMSILPKEXDKVHNPRYDGNISYPLAN 360
 DB 301 GNGVYTTISGNDGQSITQKSGIDPSEIKITLAMSILPKEXDKVHNPRYDGNISYPLAN 360
 QY 361 NGRTIYFMDOQKQSLIPASDINOAGAIYEEGNFTVSPNSQVTOGAGIHVSENSTYTWK 420
 DB 361 NGRTIYFMDOQKQSLIPASDINOAGAIYEEGNFTVSPNSQVTOGAGIHVSENSTYTWK 420
 QY 421 VNGVEHRLSKIKGTLHYQAKGENKGSISVGGKGVILLEQADDOQKQAFSEILVSGR 480
 DB 421 VNGVEHRLSKIKGTLHYQAKGENKGSISVGGKGVILLEQADDOQKQAFSEILVSGR 480
 QY 481 GTVQLNDKQPDYDKFYEGFRGRLDNGHSLFKRIQNTDEGAMYNHNTYQAAVYTT 540
 DB 481 GTVQLNDKQPDYDKFYEGFRGRLDNGHSLFKRIQNTDEGAMYNHNTYQAAVYTT 540
 QY 541 GNESTIVPNGNNINKLDYRKELIANGFGETDKNNKRLMLIYKPTTEDRTLLSGGTN 600
 DB 541 GNESTIVPNGNNINKLDYRKELIANGFGETDKNNKRLMLIYKPTTEDRTLLSGGTN 600
 QY 601 LKGDITOTKGLFEFGRTPHAYNHLNKRSMWEGIPQGLTWMDHMINRTKAEINFOIX 660
 DB 601 LKGDITOTKGLFEFGRTPHAYNHLNKRSMWEGIPQGLTWMDHMINRTKAEINFOIX 660
 QY 661 GGSAAVSRNVSIEGNMTVSNNAATPGVVPQONTICTRSDMTGLTTCQKVDLDTKVI 720
 DB 661 GGSAAVSRNVSIEGNMTVSNNAATPGVVPQONTICTRSDMTGLTTCQKVDLDTKVI 720
 QY 721 NSIRKQINSINLTNATANYKGLAKNNTLTINHSQFTLSNNTQIGNTRISDNSTA 780
 DB 721 NSIRKQINSINLTNATANYKGLAKNNTLTINHSQFTLSNNTQIGNTRISDNSTA 780
 QY 781 TVDNANLNGNVLITSAQPSLJNSHFSHOIQGKTGTTLNATWTPSDTTLQMLTLN 840
 DB 781 TVDNANLNGNVLITSAQPSLJNSHFSHOIQGKTGTTLNATWTPSDTTLQMLTLN 840
 QY 841 STITLNSAYSASSNTPRRRSILETTPTTSABERNLTVNGKLSQGTFOFTSLSPEYK 900
 DB 841 STITLNSAYSASSNTPRRRSILETTPTTSABERNLTVNGKLSQGTFOFTSLSPEYK 900
 QY 901 SDKIKLNDABEGYILSVRNGKEPTELEDTLVESKNOPLSKIKFTLENDHVDGAL 960
 DB 901 SDKIKLNDABEGYILSVRNGKEPTELEDTLVESKNOPLSKIKFTLENDHVDGAL 960
 QY 961 RYKLVKNDGEFRLLNPIKEGEJLHNDIVRAQAEFTLEAKOVBPAKTOTGEKVSRAA 1020
 DB 961 RYKLVKNDGEFRLLNPIKEGEJLHNDIVRAQAEFTLEAKOVBPAKTOTGEKVSRAA 1020
 QY 1021 RAAPDTLPDQSLNALBEAKALITAEQKSAKTKVRSKRAVSPDPLDQSLPALBA 1080
 DB 1021 RAAPDTLPDQSLNALBEAKALITAEQKSAKTKVRSKRAVSPDPLDQSLPALBA 1080
 QY 1081 LEVIDAPQSEKORLAQEBEAKOKKODLISYSNALSLSATWSMSTODELDFV 1140
 DB 1081 LEVIDAPQSEKORLAQEBEAKOKKODLISYSNALSLSATWSMSTODELDFV 1140
 QY 1141 DOAASAVMTNIAQDKRRYDSDAFRAYQOQKTLROIYQOKALNKEIGVFSHSNDTE 1200
 DB 1141 DOAASAVMTNIAQDKRRYDSDAFRAYQOQKTLROIYQOKALNKEIGVFSHSNDTE 1200
 QY 1201 DEOVKNATLTWASGAOYQWGDLOFGVNVGGISASIMAEQSRKHKAINTGVNASTY 1260
 DB 1201 DEOVKNATLTWASGAOYQWGDLOFGVNVGGISASIMAEQSRKHKAINTGVNASTY 1260
 QY 1261 QFRLGOLGOPYGVNRYFERENYSEEVRYKTPSLAENRYNAGRIVYETTPEDINISV 1320
 DB 1261 QFRLGOLGOPYGVNRYFERENYSEEVRYKTPSLAENRYNAGRIVYETTPEDINISV 1320
 QY 1321 KPEFYVYDVSNAAVQTTVNLTVLQOPFGRYMOKVGLKABLHFOISAFISKSQASOL 1380
 DB 1321 KPEFYVYDVSNAAVQTTVNLTVLQOPFGRYMOKVGLKABLHFOISAFISKSQASOL 1380
 QY 1381 GKQONVGYLGYRW 1394

DB 1381 GKQONVGYLGYRW 1394
 RESULT 2
 B81963
 IGA-specific serine endopeptidase (BC 3.4.21.72) NMA0457 [imported] - Neisseria meningitidis
 C:Species: Neisseria meningitidis
 C:Date: 05-May-2000 #sequence: revision 05-May-2000 #ext_change 02-Feb-2001
 C:Accession: B81963
 R:Parthill, J.; Achman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
 Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
 A:Reference number: A81775; M01D:2022556; PMID:10761919
 A:Accession: B81963
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1449 <PAR>
 A:Cross-references: GB:AL162753; GB:AL157959; NID:97379120; PID:CAN83754.1; PID:9737920
 C:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: iga2; NMA0457
 C:Superfamily: IGA-specific metal: endopeptidase
 C:Keywords: hydrolase; serine proteinase
 Query Match 57.4%; Score 4178; DB 2; Length 1449;
 Best Local Similarity 56.2%; Pred. No. 4.6e-196;
 Matches 816; Conservative 214; Mismatches 325; Indels 98; Gaps 17;
 QY 10 FLTACISIGVSCAMAGTYFGIDYQYRDPANENKKEFTVGAQNKIKYVNGQGLVGTSM 69
 DB 27 YLALCLSGELLQAMAGHTYFGINQYRRDPAENKKEFTVGAQNKIKYVNGQGLVGTSM 86
 QY 70 KAPMIDSVSRNRYVAALENQYVSAHNVGYTDVDPGAEQNDQREFTYKIVRNNY 129
 DB 87 KAPMIDSVSRNRYVAALENQYVSAHNVGYTDVDPGAEQNDQREFTYKIVRNNY 146
 QY 130 KCONLHVEDDYAPRLHKVTEAAPIDMTSMNGSTYSDRTKEPERYIGSGRCGEMRD 189
 DB 147 KPNSHPTNGYHMPRLHKEVTDAPFVEMTSDMKNTISDKKTYERKYSQGHRTYRD 206
 QY 190 QDKXQVAGAYHTTAGNTHNORGANGYSTLGGDYAKAGEPLPIJSGKDSGSPMT 249
 DB 207 DDHEDGLSYGAWLIGGNTHQMGANGVSLSGDYRHANDGPMPIGAAAGDSGSPMT 266
 QY 250 YDAEKQKMLNGILRENPBEKENGFOYVAKSYF-DEIPEEDLITSLYTAQNGYTTIS 308
 DB 267 YDKNNKMLNGVLOTGYPSGREGFQLRKQMEFYDDIYRGDHTYFEPBRSGHFSPT 326
 QY 329 GNDGGGSIQKSGIPSEIKITLAMS-PLKEXDKVHNPRY--DQPNYSPRLNNGE 363
 DB 327 SNNNGGIVTETNKVSNPKLKYQYTRLPDESLEITDK--EPVYAGGVNOQYPRLNNGE 384
 QY 364 TLVFMDOQKSLIPASDINOAGAIYEEGNFTVSPNSQVTOGAGIHVSENSTYTWK 423
 DB 385 NLSPFDYNGKLLSNININGAGAGIYEEGDTVSPENNETWGAQVHISEDSYTWK 444
 QY 424 VEHDRLSKIKGTLHYQAKGENKGSISVGGKGVILLEQADDOQKQAFSEILVSGR 483
 DB 445 VANDRLSKIKGTLHYQAKGENKGSISVGGKGVILLEQADDOQKQAFSEILVSGR 504
 QY 484 QLANDKQPDYDKFYEGFRGRLDNGHSLFKRIQNTDEGAMYNHNTYQAAVYTT 543
 DB 505 QLANDKQPDYDKFYEGFRGRLDNGHSLFKRIQNTDEGAMYNHNTYQAAVYTT 564
 QY 544 SYLVNNGNNINKLDYRKELIANGFGETDKNNKRLMLIYKPTTEDRTLLSGGTN 603
 DB 565 SITQSGKINIRLAYSKEIANGWFGEDOTTKTGRLNLNVQPAEDRTLLSGGTN 624
 QY 604 DITQKGLFEFGRTPHAYNHLNKRSMWEGIPQGLTWMDHMINRTKAEINFOIX 663
 DB 625 NITQNGKLFPSGRTPHAYNHLNKRSMWEGIPQGLTWMDHMINRTKAEINFOIX 684

```

QY 664 AVVSBNSSISIEGNTVSNANNAFEGVVPNOQNTICTRSDWTGLTTCQKVDLTDTKINSI 723
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 685 AVISRWAKVSGDWHLSNHAQAVFVAPBQSHITCRSDWTGLTTCVCKITTDCKVLASL 744
QY 724 PRTQINGINLTDAATANYGLAKNGVNTLNHSGFTSSNNAATQIGN----- 771
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 745 TKTDISGNVSLADHAHNLGLATLNGSLANGDTKYTSYSHATQNGMLSLVGNMAQATFN 804
QY 772 -----IRSDNSTATVDNANLNGVNLTDASQFSLKNS 804
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 805 QATMGNTSASGNASFSLSNNAONGSLTSLSDNAKANVSHSLMGNVSLADKAVFHFENS 864
QY 805 HSHQLOQDKTITVLENAWTMPSDITLQNTLNSTTTNSAY-----SASSNNT 856
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 865 RFTQGLSSKDTALHLDSEWTLPSGTSLGNLMDNATITLNSAVRDADAGAGTGVST 924
QY 857 PR---RSLETEITPTSAREHFNLTAVNGKISGGGTFOFTSLFGYKSDKILKSNDAEC 913
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 925 PRRRRRSLSLVTPPTVSRESFNLTAVNGKLNQGTFRMSLPGYSKDKILKLBSSBOT 984
QY 914 YLISVYNTGKPELTLEQTLVSSKDNQPLSDKLTFLNDHYDAGALKYKLVNDGEPL 973
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 985 YTLAVNNTGNPVSLLDQLTVEGKKNKPLSENLTFLQNEHYDAGAWKYOLIKDGEPL 1044
QY 974 HNPKEOELHNDIVRAE---QAE-----TEAKQVEPTAKTQI-GEKPVMSRAARA 1022
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1045 HNPVEOELSDKGAOKKQAKKQKQKQSLDALIAGKDAKIBSVAP-----AQ 1097
QY 1023 APEDTLPDQSLNALTEAKQELTAETOKSKAKTKVRSKRAVSDPLDQSLFALMALE 1082
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1098 AGGENV---GIMQAEKK-RVQAKDSALAKREAEPTATLP-----RARA 1144
QY 1083 VIDAPQSBKDLAQEAEKQ-RKQKDISYSNSALSSTVNSKLSVQDLDLPLPD 1141
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1145 RRLPLQ-----PQPPQPPQPPRLISRYANSGLSFRSATLANSYFNAVQDDELDFAL 1197
QY 1142 QAGSVAWNTINADKRRDSDAPAVQOKTMRIGVQKALNGRIGAVSHSRSDNTD 1201
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1198 DRNNAWTSGIRDTKHSRDFRARIQO-TDLRITQKQKSSGVVGLTSHRNTENTD 1256
QY 1202 EGVKNHATLTMMSGFAQYQWGLDQSGVNVGTGISASKMAEBSRKIKRKAINTGVASYQ 1261
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1257 DQGNASARLHAGAVFGYIGRDPDGLSTGAGFSSSLSDGCGKIRBVLHYGIQARX 1316
QY 1262 FRIGLGIQYFVGNRYFIERENYQSEYKVTSLAFPNYAGIRVDYTTPTDITSVK 1321
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1317 AGFGGGLIEPYIGATRYFQKADYREYVNIATPGLAFNRYDAGIKADYSFKPAOHISIT 1376
QY 1322 PYFVYVYDVSNANVQTTVNLVYLOQPFGRYQWQVGLKAEILHFOISAFYSKSSQSLG 1381
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1377 PYLSLSTDDAASKYKRVNTAVLMODFGKTRSBKNGVNLKGTTLISLHAAAKPOLE 1436
QY 1382 KQONVGVKLGYSR 1394
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1437 AQHSAGCKLGYRM 1449
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

RESULT 3

D81019
 adhesion and penetration protein NMB1985 (imported) - *Neisseria meningitidis* (strain MC5
 C)Species: *Neisseria meningitidis*
 C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C/Accession: D81019
 R/Reteller: H. J. Saunders, N. J. Heidelberg, J. Jeffries, A. C. Nelson, K. E. Eisen, J. A.
 Hickey, R. K. Haft, D. H. Salzberg, S. D. White, O. J. Fleischmann, R. D. Dougherty, B. A.
 ri, H. J. Qin, H. J. Vanatdevan, J. Gill, J. Scarlato, V. J. Maignani, V. J. Pizza, M.
 Science 287, 1809-1815, 2000
 A/Authors: Grandi, G.; Sun, L.; Smith, H. O.; Fraser, C. M.; Moron, E. R.; Rappelli, R.; Ve
 A/Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC5.
 A/Reference number: AB1000; M010; 2015755; EMD:10710307
 A/Accession: D81019
 A/Status: preliminary

```

A:Molecule type: DNA
A:Residues: 1-1457 <RET>
A:Cross-references: GB:AB002547; GB:AB002098; NID:G7227244; PIDN:AAF42312.1; PID:G72272
A:Experimental source: serogroup B, strain MC5
C:Genetics:
A:Gene: NMB1985
C:Superfamily: IGA-specific metalloendopeptidase

Query Match          56.2%  Score 4085;  DB 2;  Length 1457;
Best Local Similarity 55.1%  Pred. No. 1.6e-191;
Matches 806;  Conservative 219;  Mismatches 328;  Indels 110;  Gaps 19;

QY 10 FLTACISLIGVQAMAGHTYFGIDYQYRDPFAENKGTGVAQNIYVKNQGLVGSMT 69
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 27 YLALCISFLGLPQAMAGHTYFGINYYRDPFAENKGTGVAQNIYVKNQGLVGSMT 86
QY 70 KAPMIDPSVSRGVAAVLVENQYVSVYAHVGYTDVDFGAEQNNPDQHFYKLYKRNRY 129
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 87 KAPMIDPSVSRGVAAVLVDQYVSVYAHVGYTDVDFGAEQNNPDQHFYKLYKRNRY 146
QY 130 KKMNL-HVEDDYHNPRLHKFTEAPIDMTSNMNGSTYSRDTYRERYRISGRQFMEN 188
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 147 KAGTGHFPGGDHMPRLHKFTEAPIDAEVEMTSMYDGRKTIQNNNTYRORVLAGRQYRS 206
QY 189 DQDKGD-----QVAGAYHYLTAQTNHQRGANGYSYLSGQVRAKAGEYGLPIAGSKD 242
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 207 DEDPPNRRESSYHILASYSWLVGNTPAQNGSGGTVALGSEKIKESPYGFLPTGSGFGD 266
QY 243 SSGSMPTFYDAKQKULNGIIRBGNPPEKENGEPQYKSYF-DELFEDLHTSLYTAG 301
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 267 SSGSMPTFYDAKQKULNGVLTQGNPTIGKNSGFLVAKVMFVDELPADDTLSVFEBRQ 326
QY 302 NGVYTTSGNNGGASITQK--SGIPSEIK--ITLAMS-PELEKQKVRHNPYDGPVIY 355
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 327 NGKTSFNDNNGGKYNKAKHEHNSLPRLKTRTVQVPLVSLSTRAPEPTYHA-GGVNSY 385
QY 356 SPRNLNGETLYFWDQKQSLIFASDINOAGAGVTEGNE-TVSPNSQYQAGAHVSENS 445
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 386 RPLNNGENISFIDEKGELLTGNINQAGGVLFGQDPTVSPNNETQAGAYHISDS 445
QY 416 TYTWKQNGVEHRLSKTGKTLHQAQKENGKSLSYGQKYLEEQADLQNGKQKFEISIG 475
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 446 TYTWKQNGVANDLSLTGKTLHQAQKENGKSLSYGQKYLEEQADLQNGKQKFEISIG 505
QY 476 IYSGRGVQALNDKQPDNDKFPYFGRGRLDNGSLFERRLQNTGEGAMIVNHTTDA 535
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 506 IYSGRGVQALNDKQPDNDKFPYFGRGRLDNGSLFERRLQNTGEGAMIVNHTTDA 565
QY 536 NVYTTGNESIVLPNGNINKLIDYKEJAYNGVEGTDRKNGKRLIYKPTEDRLL 595
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 566 TYTTGKMDLA-TTGN-NS-DSKKEIAYNGVEGTDRKNGKRLIYKPTEDRLL 623
QY 596 SGGTYLKGDIQYKGLPFSGRPTPHAYNLNKRSEWEGIPQELTYWCHDYNTPFAE 655
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 624 SGGTYLNGNTTYNGSLFPGSGPTPHAYNLNKRSEWEGIPQELTYWCHDYNTPFAE 683
QY 656 NFOIKGSAVSRNVSIEGNTVSNANNAFEGVVPNOQNTICTRSDWTGLTTCQKVDL 715
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 684 NFOIKGSAVSRNVAKVGDHLSNHAQAVFVAPBQSHITCRSDWTGLTTCVCKIT 743
QY 716 DTKVINSIPKTOINGINLTDAATANYGLAKNGVNTLNHSGFTSSNNAATQIGN--- 771
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 744 DQXVIALTKTDISGNVSLADHAHNLGLATLNGSLANGDTKYTSYSHATQNGMLSLV 803
QY 772 -----IRSDNSTATVDNANLNGVNLTDASQFSLKNS 796
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 804 GNAQATFQATLNGNTSASGNASFSLSDNAKANVSHSLMGNVSLADKAVFHFENS 863
QY 797 AQFSKNSHSHOIQDQKGTITVLENAWTMPSDITLQNTLNSTTTNSAY----- 849
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 864 AYFHESRRTGIGSGQDIALHLDSEWTLPSGTSLGNLMDNATITLNSAVRDADAGA 923
QY 850 -SASSNNTPR-----RSLETEITPTSAREHFNLTAVNGKISGGGTFOFTSLFGYKSD 902
    ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```



```

QY 127 NNYKKONLHPYED--CYNPRHAKFVTAAPIDMTS-----NMNGSTYSRDTKYERVR 178
DB 129 N-----PHEPDYDHLPRNLKLYTEISFPLALSSVPLDNGCPRAANAYLDTRFFPR 180
QY 179 IGGGRQFWRNDQKDDVAGAYHYLTAGNTHNQGAGNGXYSLGSDVYKAGEYCPLEIAG 238
DB 181 IGGGTQVAKADGRTKRTAPAYQYLTGTPTLKVGVFQNHGLVGGSLTDQ-----PLNTYA 236
QY 239 SKDDSSPMFTYDAEKQWMLINGLIEGAPPEKNGFQ--LYRKSXYDELFERDHLTSL 296
DB 237 IADSSSPLFAFDKERNRWLAGVL---STAGDFNFENKTYLTQPEFISSTIRQYETYL 293
QY 297 YTRAGNGVYTTISGNDNGGCSITOKSGIPEELKITLAMSILPLKEXDKHNRYPGPIYS 356
DB 294 --DVLGTLTHEILWRDNGNGSTIQ--GLNERTITLPIANPSI-APQNDSEHMSSEA----- 344
QY 357 PRLNGETIYF---MDQKQSLIPASDINOAGGLYEGNGFTVSPNSQCTVQAGAHVSE 413
DB 345 -----SKTLILSSREFDNK--TLMLADNINOGAGALQFDSNFTV--VAKNHTVQAGAVIAD 396
QY 414 NSTVYTKVNGVGHORLSTKGTGTHVQANGENKSSIVGQKYLEQQADQKQKAPSE 473
DB 397 GKQVFWQVSNPKGDLSTKLAGFTILANGGGLNQDLSIGSTVLAQAAASDSKQAFNQ 456
QY 474 IGLVSGRTVQLANDKQPDIDKEFGFRGSLDLNAGSLTEKKIONTDEGAMIVNHTTQ 533
DB 457 VGLTSGRTAVLADSQQKPBENLYFSGFRGLDINGNNLAFHIHADQGAQYINHNPPQ 516
QY 534 AANVTITGNE-----STLVNPNH----- 551
DB 517 AATLTLTGNPVLSPHEVWYQWGNRPQGNAAVEYINPHNRRTDYFTIKGNGPREPR 576
QY 552 -----NPKLDYRKEIA-----YNGWGER----- 571
DB 577 LNNKSTSWQETGNRRQQAQVQAENARPDLITFGGYIGENADQTAASISKTYEAA 636
QY 572 -DKRKA-----NGRLNIYKFTTEDRTLLSGGTNLKGDITQTKSLPFPS 616
DB 637 IETRIIAMAAYGPRERYNGALNHYRPAKSTSTILLNGNNLNGVLLBGSNMLYSG 696
QY 617 RPTPHAYNLNKRMSKEGIPQCEIYWDHMINTPFAENFOIKGGSAAVNG--RNVSSIBG 675
DB 697 RPYDHAHYDQAKR-----EPVLEWMTDSFYFAAPFTLRNARLTAGRTATLDDG 746
QY 676 NMTVSNNAATGAVVQWQNTICTRSDMTLTTOOKVDLDTDTVINSIPKQIUNGSLNT 735
DB 747 DITAYDLSGIDGFTQG--KTPCEKRSYHSGSTHCIPAAVAKAENTRALPKTVQGDITL 805
QY 736 DNATAVVKGIALANGNVTLLNHSQFTLSNNATQIGNRLSDNSTATVDNANINGNVELTD 795
DB 806 DRS-----ELRL-----GKALTYG 819
QY 796 SAQPSLNSHFSHOIQSDKGTVYLE--NATWTMSDPTLLQNTLANSYITTLNSAYSSSN 854
DB 820 S-----IRAGKDTAVRMEADSNMTLSQSSHTGALTLDQAQITLTPDPFANHTH 866
QY 855 NTPRRSLFTETTPSAHRPNTLYNGCLSGQGFQF--TSLSFGYKSDKXLSNDEBQ 913
DB 867 N-----NRNPTLTVNGTLGSGFPFRLTGIVRKONAPPLKLGSDRGA 909
QY 914 YILSVNTGKEBETLEOLTIVESKQNPQSLDKETLENDHVACALRYKLVKNDGFEFL 973
DB 910 FOIHNTKGTQGEFTTESIALVSL--NPKSHARTLIONGYALGAYRITLRKNNNYSL 967
QY 974 HNPDKQELHNDLVAS-----QAERTLEAKQVEETAKTQGEPKYRSRRAAPAPDT 1028
DB 968 YNPKTAHEQIATATSAHEBRNOQAYNOQATDI-----SRQVCHSD 1009
QY 1029 PDGSLNLAEAKQAEIT--AEFOKAKAKTKKVVNSKRAVPSDELD-----OSI-----F 1075
DB 1010 ATHQALQAWONSOTELARIDSOVYVLSAQIKQ-----TDLPTGILTTRADNLCMAQGY 1061
QY 1076 ALFALEVIDAEQCS-----EKRLAQEBEAKQKQKQ----- 1108

```

```

DB 1062 SADICQVAAKADTNDLTLFETELDTYIEREWEASELDDKARQGGDAQAVETARHAYLNA 1121
QY 1109 -----LISRSNSALSELSTAYTNMLSVDELDRLFLVQQA 1143
DB 1122 LNRLSRQHSIKTVAGIRMPNLAEILISRSANPAVSEDAANTNGRQOAGRIIDLRLHTPQ 1181
QY 1144 QSAVWYNLAODKRRYSADAFRAYOQKTNDRQIGVQKALANG--RIGAVFSHSRPTDE 1202
DB 1182 QONTLWETGVCQDTHSGTHRPY--QOTTVAHIGTGTGTRLSVGLTLDEKTNRRDE 1240
QY 1203 QY-----KNEALTYMNSGPAQYQWGDLOFGVNVGIGISASQAAEBSRKIRKALINYGN 1257
DB 1241 GVSASNSNGAHL-----FVKENGALFAAADGYSNSRTRFTDYDGAAYRHRAMDAGIN 1295
QY 1258 ASYQPRQGLQIGQPYGVNRYFIREK---YQSEBRVKTSLAFNNYVNGIRHYDTFTF 1314
DB 1296 TGKIKDTG--INLRPVGIR---INRSKNRYVIDAGELNSPQIQCTWAGELRDKT--VE 1350
QY 1315 TDNISVPRYFVNVYDVV--SNANVQTTVN--LTVLQPPFG---RYWQKEVGKAEILLFQIS 1369
DB 1351 LGQATLTPASSDYHTTRONGSALSVDRTLLQQAHCGLTHTQIDNGYKS-----WNAK 1406
QY 1370 AFISKSQGSQQLGKQWVGVKGYRW 1394
DB 1407 LHAAYKQSTARRKQAGIKIGYRW 1431

```

RESULT 6

A:7023
 19A-specific metalloendopeptidase (EC 3.4.24.13) type 1 precursor - Haemophilus influenzae
 N:Alternate names: immunoglobulin A1 proteinase type 1
 C:Species: Haemophilus influenzae
 C:Date: 31-Jan-1992 #sequence_revision 12-Jun-1992 #text_change 08-Dec-2000
 C:Accession: A37023
 R:Poulsen, K.; Brandt, J.; Hjorth, J.P.; Thøgersen, H.C.; Kilian, M.
 Infect. Immun. 57, 3097-3105, 1989
 A>Title: Cloning and sequencing of the immunoglobulin A1 protease gene (iga) of Haemophilus influenzae
 A:Reference number: A37023; PMID:89379374; PMID:2506130
 A:Accession: A37023
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1541 (POU)
 A:Cross-references: GI:X64357; NID:943560; PIDD:CAA45708.1; PID:943561
 A:Experimental source: serotype b
 C:Superfamily: Iga-specific metalloendopeptidase
 C:Keywords: hydrolase, metalloprotease

Query Match 23.0%; Score 1673.5; DB 2; Length 1541;
 Best Local Similarity 28.5%; Pred. No. 1e-73;
 Matches 480; Conservative 246; Mismatches 512; Indels 445; Gaps 56;

```

QY 6 PLNLTLCISLGVQAAKHTTFTGIDYQVRCPEAKNGKFTVGAQINLVNKKQQLVG 65
DB 6 PLNLTALTVAYVLTPTETALVRDQVYQIFPDFAENKGSVGATNVLVKDKNNKDLG 65
QY 66 TSWTAA--EMIDFSVVS--RKGVALAVENQYIVSAH--NVGYTDVDFG--AEQN-----NPDQH 117
DB 66 TAPNGIMIDFSVDVVDKRIATLINQYVYGAKHSNKSQVSELHFGULNNMANNNGAAKAH 125
QY 118 RFL-----YTVKERNY-----KQDNLHPYEDYHNPRLHFTVTEAPID--MTS 160
DB 126 RVSSSEKRYFSVSKKEEPTKLNSKTVTEQOKREHRYWPRLDYFTEVALEIASSTA 185
QY 161 NMKGSITSDKTYEERVIRIGSGRCFWRDQDKQ-----VAGAYHL 203
DB 186 SSDAGTYNDQKPPAFVRLGSSGCF--TYKGDNYSLILNHEVGGNN--KLVDGAYTYG 242
QY 204 TKGANTHNRGAGKGYSLG-----GVYKAGEGPLPAGSKDGS PMFTYDAEKQK 257
DB 243 IACTPYKKNHNNKGLGFGNSKSEHSDPKGLISODLTVYVVLGDSGSPFLVYDRKGM 302
QY 258 LINGILREGNPFSGKB--NSFQLVKSYFDEIFERDLHTSL-----YTRAGNG--VTTISG 309

```



```

Db      303 LFLASYPFAGNFKS... 362
Qy      310 NDNGQSGITOKSGIPSEIKITL... 363
Db      363 GE-----KSLNVDLAD-... 366
Qy      370 QKQSSLIIFASDINGAGLIFEGNF... 428
Db      387 EGSGGTLNNNDIDQAGGIFPEGD... 446
Qy      429 LSKIGKTLVACAKGKNGSISVGD... 488
Db      447 LAKIGKTLVBEQDNGSKLKVGD... 505
Qy      489 KQPDTCYFEGFEGSLDINGSLT... 548
Db      506 KQVDPNSITYFPGFGLDINGSL... 565
Qy      549 N----- 549
Db      566 KTTTPYIDAPDENPYAFRIK... 625
Qy      550 -----GNNINKLDYRKEI... 597
Db      626 KLYMGKTSDEAKNNVNNHIN... 684
Qy      598 GTNLKSDITQTKGKLIFPSGR... 647
Db      685 GTNLGDLTVEKGTLIFLSGR... 726
Qy      648 INRTEKAKNFOIKGSAVVS-... 706
Db      737 INRPFKATMTAVTGNASLIS... 794
Qy      707 TTCQKVDLDTKYNSIPETQ... 766
Db      795 VCTCTDKLSD-KALNSF... 845
Qy      767 TDIGKTLSDNSTATYD... 826
Db      846 -----VRLTENS-... 871
Qy      827 MPODTTLQNTLNNSTTL... 886
Db      872 -----IHANSA--... 896
Qy      887 QCTPQPTSSLFYKSDKL... 946
Db      897 NGSFYLLTDLNKGQKVV... 953
Qy      947 KFTLENDVDAKLRK... 987
Db      954 NWSLVGNTVDLGA... 1011
Qy      988 -----RAEQAKERTLE... 1004
Db      1012 SNNEELARVDEAVPP... 1071
Qy      1005 -----AKTQYGE-... 1020
Db      1072 AKSNYKANTQTNNEVA... 1131
Qy      1021 RAAPFDTLPROS-----... 1071
Db      1132 KQBSGTVQQAAPAREND... 1185
Qy      1072 QSLFALBALEVIDAFOQ-... 1107
Db      1136 ST--TVWNGSVVENENT... 1243
Qy      1108 -----DLISRYNSALS... 1154

```

```

Db      1244 RSTVALCDLSTNTINMAV... 1303
Qy      1155 KRTYSDAFRAYQOQNTN... 1213
Db      1304 MKVYSSSYRRFSSKSTQ-... 1361
Qy      1214 SGPAQY-----QWEDL... 1269
Db      1362 NRYSKTYADNM--YIG... 1418
Qy      1270 QPYFGVNRYPFERENTY... 1329
Db      1419 TPIVGRYSYNSNAD... 1476
Qy      1330 DVSNAVQTTNLTIV... 1389
Db      1477 DAKQSGSKNTNGV... 1536
Qy      1390 LGY 1392
Db      1537 LSF 1539

RESULT 7
B41859
IGR-specific metalloendopeptidase (EC 3.4.24.13) type 1 precursor - Haemophilus influenzae
C:Species: Haemophilus influenzae
A:Variety: Strain HK393
C:Date: 04-Mar-1993 #sequence, revision 18-Nov-1994 #ext_change 08-Dec-2000
C:Accession: B41859
J:Paulsen, K.; Reinholdt, J.; Kilian, M.
A:Title: A comparative genetic study of serologically distinct Haemophilus influenzae ty
A:Reference number: M41859; MUID:92234949; PMID:1373717
A:Accession: B41859
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-1545 <POU>
A:Cross-references: GB:M67490; NID:G148908; PIDN:AAA24967.1; PID:G148909
A:Experimental source: strain HK393
A:Note: sequence extracted from NCBI backbone (NCBI:97283)
C:Superfamily: IGR-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase

Query Match      22.9%; Score 1662.5; DB 2; Length 1545;
Best Local Similarity 28.6%; Pred. No. 3.6e-73;
Matches 482; Conservative 246; Mismatches 511; Indels 447; Gaps 59;

Qy      6 FRNLFLTACISLGV... 65
Db      6 FKLNFALVAVALT... 65
Qy      66 TSWTKA-PMIDP... 117
Db      66 NVLPNGIPMIDP... 125
Qy      118 RFT-----YKIVK... 160
Db      126 RUVSSENRITYVER... 185
Qy      161 MNNGSYSDRYT... 209
Db      186 SSTAGYNNKQK... 243
Qy      210 NQSGANGYSY-... 254
Db      244 TYGIAGTPEVNH... 303
Qy      255 QKMLINGILBEGN... 306
Db      304 GKMLFLGSYD... 363
Qy      307 ISGNDGQSGITOKSGIPSEIKITL... 366

```

```

Db      364 ITGGE-----KSLANDLAD-----OKDP-----NHGKSVT 389
Qy      367 FMOQKGSLLPASPINOAGAGLYFEGNFTVSPNS--KOTWQAGIHVSENSTVWTKNGVE 425
Db      390 F--BSSGILLNNNDQAGGLFFBSGDIYEVGISTDNITWKGAGVAVAGKTVTKWKNQ 447
Qy      426 HDRLSKIGKGLTAYOAKGENKGSISVGDGKVLLEQADDDQNKAPSEHIGVSGRGVQJL 485
Db      448 YDLATIKIGKGLTIVEGTDKSGSLKAGDGTILKQNTSSG--QHPASVGVISGRSLVIL 506
Qy      486 NDDKQPTDDKTFYFGRGGRJDLNGSLTPKRIQNTDGMATYNNHTTAAAVYITIGNESTI 545
Db      507 NDDKQVDPNSIYFGRGGRJDLNGSLTFEDIRNIDEGALVHSTSGKSTVYITGDNDLI 566
Qy      546 VLENG----- 550
Db      567 TDRNNSIYTYKPLLEDNPATIKQIKYQULYFNEENKTYTAKADASIRSEFPQNRGS 626
Qy      551 -----NNINKLDYRKELAYNGMFGETDKNKNHGRNLNLYKPTTERLIL 594
Db      627 NNSMLYMGTEKADQKAMNHINNEBNGFNGYFGE--BEKKNNGNLNVTFKGKSEQRRELI 685
Qy      595 LSGSTLKCDITQIKGKLFSSGRPTPHAYNHLNKGSEAGI-----PQGEIYWD 644
Db      686 LTGGLTNGDLNVQGLFLTSGRPTPHA-----RDLAGISSTKSDSHPSENNVEYVE 737
Qy      645 HDNINTEFAENFOIKGSGSAVVS--RVVSSIEGNNTVNNNMTFVGVEPQOQNTCTCSPM 703
Db      738 DDMINERFQATINVTNNATLYSGRNVESITSNTAANNKAKHITG--KAGDTICVARSIDY 795
Qy      704 TGLITQCKVDLTDEKVINSLPKTOINGSINTLDNATYVAGLAKINGVTLTNHSCPTLS 763
Db      796 TGYVCTCTDKSD--KALNSNPNTNLGNVNLTESANP--VLGKANLFGTIOGRNGS--- 849
Qy      764 NNATQGNTRISDNSTATVMDNANLNGNHLTSAQFSLKNSHSHOIGDGGTIVTLENA 823
Db      850 -----VALTENS--HMHILTEN--SDVQDLNLAGH----- 875
Qy      824 TWMPSPDTTLQKLTNNSTLINSAYSASNNTPRRSRLETTPTTSAEHRFTLTVNGK 883
Db      876 -----IHLNSA--DNSNNVT-----KYNULIYN-S 897
Qy      884 LSGGSEFOPTSLFQVSDYDKLKLSNDAEGDYLISVRYNGKEPPTLEQTLVSESXONQPLS 943
Db      898 LSGGSEFYITDLSKQGDQVWVTGATGNFTLQVADTGEPRH--NELTLDGASQAQ--R 954
Qy      944 DKLKFTLBNDDYDAGALRYKLYKNDGEPRLNPIKEQELNDLY----- 987
Db      955 DMLNVSIVGNVTDLGAMKYKLRVNGRYDLNVP--EVEKQQTVDTTNITTPNNICADYV 1012
Qy      988 -----RABQERTLEAKOVEPT----- 1004
Db      1013 SVPSNNEELARDEPRPPAPATPSETTETVLAENSKQESKTYERKQDAJETTAQUREV 1072
Qy      1005 -----AKTQGE-----PKYRS 1017
Db      1073 AKEAKSVKANTCTNEVAGSGSEKETEOTTEKETATVEKEKAKVEKETEQTQEPKYSQ 1132
Qy      1018 RAARAAAPROTLDQS-----LNLALAK--QALITPQSKAKTKKVKNSKAAVSD 1068
Db      1133 VSPQQSETEVPQAEPRNDPTVNIKEPQCTNITLADTEQPAKETS-----SVEQ 1186
Qy      1069 ILDDSLFALALAVTAPQO-----SEKRLAQEAEKQKQK----- 1107
Db      1187 VTST--TWYTGNSVHEPENTPATTOPTVNSSESNKPKNRHRSVRSVPHNVPATIS 1244
Qy      1108 -----DLISYSNALSLSL--TWNSMLSVDQELDLFLV--DQAGAVWNTI 1151
Db      1245 SNDSFTVALCDLITSTNTAVVSDAKAQFVALNIVGKAVSOHLSOLEMNEBQYVAVSN 1304
Qy      1152 AQDRRYDSAFRAVQOQKNTLRQIGVQKALNG--RIGAVYSHSBDNTFEDQVYKHAFL 1210

```

```

Db      1305 TSNANKTSSQYRPRSSSTQ--TQGMQDTISNNQQLGQYVTVYRNSNPFQATSKA--TL 1362
Qy      1211 TWMSGFAQY-----QWGDLOFQVNVGTGISASKAEQSRKIRHKAIVYNASVYQFRLQ 1266
Db      1363 AQVNVFQKTYADNMH---YLGDIGYCKPQSKLQNNNAKARHTAQFGLTGKAPNLGN 1419
Qy      1267 LGIDPYFQVNRVYFIERENYQSEEVAVTSPSLAENRNNGIRVDYTFPTDMISVXYFV 1326
Db      1420 FGITPIYGVARSYLSNADPFDQARIKVNPISVKTAFQAVDLSTYV--HLGEFSVTPILISA 1478
Qy      1327 NVVDYSNANQCTVMNLTLYQOPFRWQKEVGLKXELHLPQISAFISKSGSGLQKQNV 1386
Db      1479 RT-DNMGSSKINVNGVDYFNVNENQQYNAIGLKIKYHNVKLSLGGTLKQAQAEKOKTA 1537
Qy      1387 GWKLG 1392
Db      1538 EKLKSP 1543

RESULT 8
S61314
IGA-specific metalloendopeptidase (EC 3.4.24.13) homolog - Neisseria meningitidis
C:Species: Neisseria meningitidis
A:Variety: Hfl3
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 08-Dec-2000
C:Accession: S61314
R:Lombolt, H.; Poulsen, K.; Mogens, K.
Mol. Microbiol. 15, 495-506, 1995
A:Title: Comparative characterization of the iga gene encoding Iga1 protease in Neisseria
A:Reference number: S61314; PMID:95502861; PMID:7783620
A:Accession: S61314
A:Status: preliminary
A:Stature: preliminary
A:Molecule type: DNA
A:Residues: 1-1561 <LOM>
A:Cross-references: EMBL:X62474; NID:q732873; PDB:QAA57857.1; PDB:q732874
C:Superfamily: IGA-specific metalloendopeptidase
C:Keywords: hydrolase; metalloprotease

Query Match      22.7%; Score 1654.5; DB 2; Length 1561;
Best Local Similarity 28.8%; Pred. No. 8.9e-73;
Matches 494; Conservative 239; Mismatches 508; Indels 473; Gaps 55;

Qy      1 MKKTVRLNFIACISL--GYSCAMAGHYFGIDYQYRPAEKKGKTVNGCINXYN 58
Db      1 MKTKKFKINALSISIFLAVALPYSEALVVDVQIFRDAENKKGKFPVGAITLSDVKN 60
Qy      59 KQGLVGTSTKAPMIDFVSGRN--GVNALVENQYIVSYAHNV-----GYTLY 105
Db      61 KQGNIGNALSNNPMDIDFSVADYVKGATVVDQYAVASVYKAKAEHTFYGGYNGHNV 120
Qy      106 DFGAGNNPDGKREFTYKIVKRNKY-----KNLHPEYEDYENPRLAKFVTEADIDWT 159
Db      121 ---ADKEN-----EVRVVEQNNYEPHKAWSASNLRL--DYNNARFNKFEVETVADIPT 170
Qy      160 SNMNG-STYSRTYTPRRYRIGSSQ-----FWNDDKQ--DQAAKAYITLAK-- 206
Db      171 DAGGADITKQKNNFSSFVRVAGGRQLYEEKGAYHDEGNEKGYDLRDSQAYRYALDGRP 230
Qy      207 -----NTHNQAGN--GYSYLGDYVAKAGEVGPLPIAGSKGSGSPMFIDAEKQ 255
Db      231 YKDINIDQNTTBGLIFGNHNTYSAELKQALSODALTNVGVUGDSSGPLRAFPKQKN 290
Qy      256 KWLINGILRENGPEKEN--GFOLVRKSYDEFERDLHTSLYTRANGVYITSGNDNG 313
Db      291 QWFLGTVDYVAGY--GKSSQWQMNITKKFADKIKORD-----NAQ---TLKXGVGH 338
Qy      314 QSGITQSGITSEIKITIANMSPLKEKDKYHNPRDGNISYPRLNKSETLVPMQKQSG 373
Db      339 HMKCTGTSIHGSTAVLAN---NEED-----ANKQGVTF--ENN 375
Qy      374 SLIFASDINQAGAGLYEAGFTV--SPNSQWQAGIHVSENSTVWTKVNGVSHDLASKI 432
Db      376 TLVLDQNTINQAGSLFFKPDYTVKANNNDITWLAGAGIDVADQKTVWQVKNRPNKDRILAKI 435

```

```

433 GKGTLHYQAKGENKGISVGPQKVLLEQADQCKOAFSEIGLVSRTGLDNDQCFD 492
436 GKGTLKINGVQOGLKNGDGLVILNOKADSNOKVSAFSGVSGRTGLVLSNSQIN 495
493 TDKRYPGFRGRLDNGSLTPKRIQONTDEGAMVNNHTTQAAVNTTGNESIVLENG-- 550
496 PDNIYFPRGRLDNGNDLTFEHIHNVDEGARIYNNHTTSLASITLTKSLITNPJSL 555
551 ----NNINKLDT-----RKET----- 562
556 VHSIQNDYDEDDYSYYPRPRPIPOGKOLYYKRYVALKSGGSVNAPEPENGQTEENDW 615
563 -----AYNGMFGEDQKNGKNGRLNIYKPTTREDRTLLSGG 598
616 IIMOSTEEZAKKXMMHKKHKNORISGSGFGEHNGKNGHMLNPNKGSQAQNFRLTGG 675
599 TNLKGD-TTOTKGLPSPGRTPPHAYNLANK-----FWSMEBQIDQSEIIVWDHWTNKT 651
676 TNLNGKISVTQGNVLLSGRTPPHARDPVNKKSSARXDAFSK-----NNEVFEEDWIMRT 720
652 FKAEHFOIKGSGSAVVS--RNVSSIEGAMTYSNNAAATFGVVPNQONTITRSMTGTLTQC 710
731 FKAEIIVNQSASFSSGGSVSNITANITATIDNAKVNLCY--KNGDEVCEVSDYTGYYTCN 788
711 KVDLTPDKVINSIPKTOINGSINTLDNATANYKGLAKXNGVTLTNHSGPTLSNNAQIG 770
789 TDNISD-KALNSFDATQINNNVNNQNA-ALVIGQAALMG----- 826
771 NIRLSDNSTATVNNALNGNWHLTDSAQSLKNSHSHQIGDKGTTVTL-ENATWTMPS 829
827 -----QIQGQNSVSINLSHSMHLTS 848
830 DTTIQNTLNKSTITLNSAYSASNNTPRRBSLETFTPTFAHREPTLVNKGKLSQGT 889
849 DQVYNISLADSEIHILNMASSDAQSAN-----KXHTKIKI-HLSGNHG 889
890 FQETSSLFGYKSPDKLSDNABGDYILSVANTGKPEPTLEQTLLVESKNQPLSDPKFT 949
896 FHYLTHLAKMLGKVLKVESASGHYQIHDQTKGPNQ-BELNLPASSVQDRS-RLSVS 947
950 LENDHVDAGALRYKVLKXNDGSEFRLNDP----- 976
948 LANNHVDIGALRYTITKTEGTRILNYPENRRKVPAPSPATYASQAQTSQIAKXPQ 1007
977 -----IKXOEHLNLUV-----RABQA 992
1008 NTVAPSPQANQAEAKRQQAQAEQYKQQAERKVAHQAEAKRQODALARQQAEOB 1067
993 ERTLEAKOVEPTAKTQOTGEPKYSRRRAARAAPDTLPPQSILNMLE-----AKQAEI 1044
1066 RQRLEAKRQAALAKQKAAEBAERKRAELAQKAAAEAKRQAALAEQAERKAAEL 1127
1045 -----TAEYOKSAKTK-----KVSXKAYVSDPL 1070
1128 AAKQAKTERKAAELAEQAAEAREAAELAKQAAEERQAAQSOPTRRNRRAIPPELSS 1187
1072 DQSLFEL-----EALFVIDAPQOSE-----KDLABEAEKQK 1105
1188 DATTTRALPRIARNSNDPSDYBEIPLDALDEDEVSVDTSQKPOUNTELHEKVTYVSL 1247
1106 -----QKDLISRYNSNLSLTSATVNSMLSVODELD----- 1136
1248 QPRAAQAPRAQAQAQQAQAVADAVSTNSNLSLADMAASQSTL-----LDTGASLTRH 1302
1137 ---FLFTDQAGSAVNTNIAQDKRYDSDARAYQQAQKTNIAQIVQAKALA-NGKIGAFHS 1192
1303 IAKQSRPAQAEKNSVNNNSNGYGRDYASAQRRTSSKRTQ-QIQISIRISSEMMQIGSYLT 1361
1199 HSRSDNTPDEQVNNKATLTMMSGPAQYQWCDLQF-GVAVGTGTSASKMAEEOSSKILHKA 1251
1362 YSDSHITFD-QASGAKTVPQANLYGKTYLNDAMVYASDIGAGSLRSLSQTKQAKANWTS 1420

```

```

QY 1252 INGVNAYGSEFRLQGLQIOPFVGARFIERENYQSEHVRVTPSLAFNNYAGIRVDT 1311
DB 1421 IQNGTLNGLNTAKINQFPIVPSAGIRYSLSADYKLGANSVVSMSKTLTGLDFAFR 1480
QY 1312 FTFPTNISVKPY-----FVNYV-----VDV-SNAVQTTNLTVLQOPFGRYQKVELKAE 1362
DB 1481 F-KVGNILYTKELLSAFAFANYGKGVNNGVNGVSFNYRADN-----QQQVSAGNA 1527
QY 1363 ILHFQSAFISKS--QOSQUGKQGVNKGKLYRM 1394
DB 1528 LITRVNLTNVNGSITKQKQLEKQSGQIKIQIRP 1561

RESULT 9
A26033
Iga-specific metalloendopeptidase (EC 3.4.24.13) precursor - Neisseria gonorrhoeae (stra
N:Allertate names: Iga protease; immunoglobulin A1 proteinase
C:Species: Neisseria gonorrhoeae
A:Variety: strain MS11
C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 08-Dec-200C
C:Accession: A26039; S09386
A:Accession: A26039
A:Molecule type: DNA
A:Residues: 1-1532 <POH>
A:Cross-references: GB:X04835; NID:g44865; PTDN:CAA28538.1; PTD:g44865
A:Note: The authors translated the codon AAG for residue 668 as Asn
R:Halter, R.; Pohlner, J.; Meyer, T.F.
EMO J. 8, 2737-2744, 1989
A:Title: Mosaic-like organization of Iga protease genes in Neisseria gonorrhoeae generat
A:Reference number: S09386; MUID:90060036; PMID:2511009
A:Accession: S09386
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 281-325, 'N', 327-337, 'N', 339-427, 'M', 429-531, 'N', 533-615, 'V', 617-631, 'N', 633-
A:Experimental source: strain MS11
C:Genetics:
A:Gene: Iga
C:Superfamily: Iga-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase; transmembrane protein
F:1-27/Domain: signal sequence #status predicted <S1G>
F:28-1532/Product: immunoglobulin A1 proteinase #status predicted <MAT>
F:986-987/Cleavage site: Pro-Ser (autolytic) #status predicted
F:1018-1019/Cleavage site: Pro-Ser (autolytic) #status predicted
F:1121-1122/Cleavage site: Pro-Ala (autolytic) #status predicted

Query Match 22.2%; Score 1617; DB 2; Length 1532;
Best Local Similarity 28.4%; Pred. No. 5.9e-71;
Matches 481; Conservative 237; Mismatches 517; Indels 456; Gaps 53;

```

256 KMLINGILREBNPEKREN--GEOLVRSYEDIEFERDLHTSLATYRAGNVTITSENDG 313
 291 QWFLCTDYDYGAGY-GRKSGWEMNITKKEFPADKIKQND-----NAG-----TVKNGEH 338
 314 QGSITQKSGIPSEIKITILANMSLPLKEXDKYHNERDYDGPITYSPILNAGETLYPMQOKG 373
 339 HMKITGINSHLGSTAVLAN-----NEGQ-----ANNQGVTFED--NG 375
 374 SLIFASDINOAGGLYFEGNFTY--SPNSNQWQAGIHTSBSNITTYWKNVNGVEHDLSTKI 432
 376 TLVLNQNINOAGAGLFFKGDYTYKANNDDITMLGAGIDVADCKVWQVKNPGBDLAKI 435
 433 GKGLHVAQKBNKSGISVGDQKVLIEQADQDQKQAFSEIGLVSGAGTQVNDKQPD 492
 436 GKGLHNGTQVNOGQGLKVGDTVLNQPALADKQVDSQVGIYSRGLVJLASSNQIN 495
 493 TDKFYFGRGRLDNGHSLTFPKRIQNTDEGAMTVNNTTOANVTITNGESITLPM--- 549
 496 PDLVYFGRGRLDNGHSLTFPKRIQNTDEGAMTVNNTTOANVTITNGESITLPM--- 555
 550 -----GNINIK----- 555
 556 VHSIQXDYDEDDYSYRRPRRPPQKDLYYKANYRYALKSGRLNAPENGVANNDW 615
 556 --LDYRKEIA-----YNGFGETDKNKNKGRRLNLYKPTTEDRLTLSSG 598
 616 IPKGTQGEBAKKNMKNKNNRIRGPGFEPBENGKCHALNIPNKSQKGRFILLTG 675
 599 TNLKGDITQTKGLFSGRPTPRAYVHNLKWSMEG--IPOGETWMDHMINRTKAEN 656
 676 ANNGKISVQGVNLISGRPTPRARDFVNSAKADHSPKNNVEYFEDDWIRTEKAR 735
 657 PQKGSASVVS--RVVSSIEBNVTVNNNAATFSGVTPVQOVTICTSDPTGILTQKVDLT 715
 736 IAVNGASAFSSGRVNSDITANTATDANKVNLG--KNQDEVCVSDTGVYVCTNGNIS 793
 716 DTGVINSIPKTOINGSINTLTDNATAVNGKALKNAGVTLTNHSGFTLSNNATQIGNIRLS 775
 794 D-KALNSPATRIRKNNVNLNQA--ALVTCALNG----- 826
 776 DNSTATVDNANLNGVNLTDQAQFSLKNSHPSHQIQGDKGTVTL--ENATVMSDTLQ 834
 827 -----KIQCGMSVSGSLNHSKMLHGLDGSQVH 853
 835 NLTNLNSTITLNSAYSSANNTPRRRSLETCETTSKHEKNTLVNKNKSGCGTQFIS 894
 854 NLSIAOSHILNMSDAQSAN-----KHTIKIN--HLSGNHPTVLT 894
 895 SLFGYSKDLKLSNDAGDYTLISVNTGKEPETLEQLTVESQNPQLSDLKFTLENDH 954
 895 DLAKTGDKVLVKSASGHQQLHVQKKTGERNQ--EGDLDFDASSVQDHS--RLPVSLANHY 952
 955 VDAQALRYKLVKNDGEPRLNP----- 976
 953 VDIGALRYTIKTENGITRLVNPYAGNGRPFVPAPSPAAVNTASQAKATQTDGAQIAKPN 1012
 977 -----IKQELHNDLVRAEQAEHTLEAKQVETAKTQGE----- 1011
 1013 IVVAPSPQANQEBALRQAKAEQVKAQDAEAEKVARQKQEBKRRRAIAIAQOEBAK 1072
 1012 -----PVRSRRAPAAAPP---TLPD--- 1030
 1073 KAMELAKQKAEARKKRELARQKAEASHQANKPRRRRRRLIPRPAPVFLDDVDA 1132
 1031 -----OSLIN-----ALBAQAEALTFQTKSKATKVKSKAVVS 1066
 1133 KDNSESSIGNLAPVIFPMGRELINDYEBIPIEB--EDAEAEERQATOFHSKSRRAISS 1192
 1067 DPLDPSLFALEALVLDAPQ--SEKQMLQCEAEKQK-- 1106
 1193 EBSDE-----DASESVSTSDGHPODTELEHEVETLAGLQRAAQPTQAAQ 1240
 1107 KDLISRYNSALSSELSATVNSMLSVQDEL--RLFVDAQSAVMTNIAQDK 1155

1241 ADAVSTVNSALSADMASTQSL-----LDTGAVYLRHIAQKSLAABKQSVMSNNGYG 1295
 1156 RYSDSFRAYQOQKTNROIGVQKALA--NGRTGVFHSBSNDTDEQVKNHATLWMS 1214
 1296 RYASAQVRRFSKRO--TOIGIDRSIENMOIGGLVYSOSHTPD--QAGKNTVQAN 1353
 1215 GPAYQWGDQF--GVNVGNGISAKMAEGRKIRKKAIVNGVNASYORLQGLIOPEF 1273
 1354 LYKGYTLNDMRYAGDITAGSLRSLRLOTOQKANKFRSTIOGILTANLTIKIQPELVSA 1413
 1414 GYRSRLSSADYLGSDDSVYVSMAVKTLAGIDPARYF--KVGLTYKPLLSAAYFANYG 1472
 1329 ---VDVSNANQVTNVLVLOPPGRGWQKWLKAEHLHQAIFSSS--QSSQGLKO 1383
 1473 KGVNVGKSEFAYKADN-----QOQVSAVALLNRYVTLNNGSITKQKQLEK 1521
 1384 QNVGKLGVRM 1394
 1522 KSGQIKQIRF 1532

RESULT 10
 H64106
 IGA-specific metalloendopeptidase (EC 3.4.24.13) type 1 - Haemophilus influenzae [strain
 MAltexmate names: immunoglobulin A1 proteinaase type 1
 CSpecies: Haemophilus influenzae
 A.Variety: strain Rd KW20
 C.Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Dec-2003
 C.Accession: H64106; A41500
 R.Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage,
 J.; Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.J.; Glodek, A.; Kelley, J.M.; Weidman,
 D.M.; Brandon, R.C.; FINE, H.D.; Fritchman, J.L.; Putnam, J.L.; Geoghegan, N.S.M.
 Science 269, 496-512, 1995
 A.Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, E.O.; Venter,
 A>Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A.Reference number: A64000; M01D:95350630; PMID:7542800
 A.Accession: H64106
 A.Status: nucleic acid sequence not shown; translation not shown
 A.Molecule type: DNA
 A.Residues: 1-1694 <TIGR>
 A.Cross-references: GB:U32779; GB:A42023; NID:91574009; PIDN:AAC2651.1; P2D:91574019;
 R.Gerndy, F.J.; Plaut, A.G.; Wright, A.
 Infect. Immun. 58, 320-331, 1990
 A>Title: Localization of the cleavage site specificity determinant of Haemophilus influ
 A.Reference number: A41500; M01D:90129281; PMID:2105270
 A.Accession: A41500
 A.Molecule type: DNA
 A.Residues: 1-377 <Gen>
 A.Cross-references: GB:X59800
 A.Experimental source: strain Rd KW20
 A.Note: the authors translated the codon TGG for residue 319 as Thr
 C.Function:
 A.Description: this proteinase is classified as type 1 because it cleaves at a proline-
 C.Superfamily: IGA-specific metalloendopeptidase
 C.Keywords: hydrolase; metalloprotease

Query Match 22.1%; Score 1609; DB 2; Length 1694;
 Best Local Similarity 26.5%; Pred. No. 1.7e-70;
 Matches 479; Conservative 251; Mismatches 536; Indels 542; Gaps 52;

6 PLNLTACISLIGSQAQAGHTYTGIDYQYRPFENKGFYTGAGNITVYKQGGQVNG 65
 6 PLNLTALTYAVALTYTBAIVRDVDYQIFPDPAENKGRFSVGAIVNEVRDKNNHSLG 65
 66 TSMTRG--PMIDFQVVS--RNGVVA--VENQYTVYAH--ANGYTDVDFG--AEQN---NPDQ 117
 66 NTLPLNGIPMDIDFVVDVVDKRIATILNQLYVYVGVKHSNGVSEHLFGNLNKNMNGAKASH 125
 118 RFT-----YKIVKNNY-----KQNLHPYEDDYNRPLHFTVTEAPID--KTS 160

```

Db      126 RVSSSENRKPFSEKAEVPTKNGKAVTTEDQKREDDYMPKDKVTEVAPLEASTA 185
Qy      161 NMNGSIYSDRTKYPERVIRIGSGQRFWRNDOKDQ-----VAGAYHYL 203
Db      186 SSGAGTYNDONKTPAVRLGSGSQP---YKKGDVNSLLINHHVGGNNLKLVGDAYYG 242
Qy      204 TPAKTHORAGANGSYLG-----GVRAAGYGPLPLASSKDDSGSPMETIYDAEKYM 257
Db      243 IAGTPYKVNENNGLLGFENSKHEHSDPKGILLQDPLTNVAVLSDSGSPFLFYDEKGM 302
Qy      258 LINGLIEBGNPFEGKE--NGPOLVRSYFDEIPEKDLHTGLYTRAGNVTTISGNDGGS 316
Db      303 LELGSDYPAAGTCKKSGWEMNITYKEPAKTVLDKDTAGSL-----TGS 345
Qy      317 ITQSGIPISEIKITLANNSLPKKEKDVNPRYO--GPNISYRLNGETLYMDQKQSL 375
Db      346 NTQVNMPTGKTSTVSNQS-----ESLNDVLDPSQDPTDKKNNHCKSVTL--RSGGTL 397
Qy      376 TPASDINOGAGLXFEKGNFTYSPNSNO--TMOGAGLHSEBNTVMKNGVEHDLKTKG 434
Db      398 TLNNNDQAGAGLFFEGDYEKGTSDSTTKGAGSVADGKVTVMKHNPKSDRLAKTG 457
Qy      435 GTLHVQAKGKNGKSIISVGDKVTILEQQADDQKQAFSEIGLVSGRTYQLNDHQEPDT 494
Db      458 GTLYVEKGENKSLKVGDTYLLKQADANNKVFAGVGI VSGRSTYVINDKQVDN 517
Qy      495 KYEFGPRGLDNGSLTPYKIONTEGAMVNTTQANNVTITNSIYLEN----- 549
Db      518 SYEFGPRGLDNGSLTPYKIONTEGAMVNTTQANNVTITNSIYLEN----- 577
Qy      550 ----- 549
Db      578 NIDAEEDNPEAFRIKQGGQLYNLENTYVALRKAGSTSEI.PKNGSESNMLYNGK 637
Qy      550 -----GNINKLAEKEIAYNGFGETDKKNGKGLIMLYKPTEDRTLLSGGTNGK 623
Db      638 TSDAKXNMWHLNHERMNGNGYFGE--BEGKNGKGLAVTKGSEGNRLILGGLTNG 696
Qy      604 DITQTKKLPSPGRPTPHAYN--HLNKSEKEGIPQSHYMDHMTNRTPAENFQJGK 661
Db      697 DTKVEKGTLEFSGRPTPHARDIAGISTKDKQHPAENRVEEDWLNRFKATYINTN 756
Qy      662 GSAVYS--RVNSTEGWTVSNNAATFGVVPNQNTICRSDDMTGLTQCKVDLTETKYI 720
Db      757 NATLYSGRVAANTITSTTASDAKVNIGY--KAGDVCYVRSDYTYTCTTMDKSD--KAL 813
Qy      721 NSIPTQINGSLMTNATANYKGLAKNGAVTLTNHSGFTLSNNATQIGNIRLSDNSTA 780
Db      814 NSFNATVSGNVNLSGNANF--VLGKAMLFGTISGTNSQ----- 851
Qy      781 TYDNANLNGNVHTDQAQSLKNSHFHQIGKKTYYTLENATYTMPSFTLLQVLTLAN 840
Db      852 -----VRLT-----BNSEHMLTGSQSNVQNLNDK 875
Qy      841 STITLSAYSASNNTPRRBSLETPTPTSAEHRPNTLVNKGLSGGGTGCFSSLPFYX 900
Db      876 GHILHNAONDAKVT-----YNTLTVN--SLSGNSSTYLTLSKQ 916
Qy      901 SDKLTSLNDAGGYILSVNTGKEPTELEQLTVESKONPLSDKLFLLIENDHAGAL 960
Db      917 GKXVYVTSKANGFTLQVADKNGE--TKNLTLLFDA--SNATKNMLVSLVNGYTVDLGAW 973
Qy      961 RYKLVNDGEERLHNAIKQELANDV----- 987
Db      974 KYKLVNNGRYDLYN--EVEKXQYVDYTNITTPNNTIQADVPSPNNEELARVETVP 1011
Qy      988 -----RAQAPRTLE-----ARQVETAK--TQVGE-- 1011
Db      1032 PAPAETPSETTEVAENSKQESKVEKQEDATETTAQNGVALEAEPKANTQTNVYA 1091
Qy      1012 -----PKYRSRRARAAP-----DTL 1028

```

```

Db      1092 QSGSETTEGTTEIKETAKVEKEKAKVEKDEIQEAPQMASETSPKQAPKVESTDK 1151
Qy      1029 PQOSLINAL-----EAQAPLTFETQSKA----- 1053
Db      1152 VEETQVQAPQOSTTVAAMAEATSPRSKAPSETQSEKTNAPVTPVSKXQNTENTDQ 1211
Qy      1054 -----KTKVRSKR-----AVESDPLDQ----- 1072
Db      1212 TEREKTAKEITEKQEPPOVASQSPKQESYTPQAPVLESENPTVNAEVEQAQLOT 1271
Qy      1073 -----SLFALAEVVIDAPOQ----- 1089
Db      1272 QTSATYSTQAPAPENS.LNTGSAATITETAFKSDKQDETASTEDASGHKANTVADNSVA 1331
Qy      1090 -----SEKRLAQE-----AEKQKQK----- 1107
Db      1332 NSESDPKSRRRRSIQPETSABETTAETASTEDTTADSKRSKPNRSRVSREPTV 1391
Qy      1108 -----DLISYSNALSLSA-----TVNSMLSVODELRLFY--DQASAWT 1149
Db      1392 TNSDSSTVALRDLTSTNTNNAVSDMAKAQFVALNKGAVSQHLSQLENNNEQYVWV 1451
Qy      1150 NIAQDRRYDSAPRAVQOKTNKQIGVQALANG--RIGAVFSHSRSDNTPEQVKHA 1208
Db      1452 SNTSMENYSSQVRRFSKSTQ--TQLGMDQTISSNVOLGVFYTVRNSNPD--KASRN 1509
Qy      1209 TLTMSGFAQY-----QMGDLQFQVNVGTGISASQMAEBSRKIKHKA.LNYVNASVQFRL 1264
Db      1510 TLQVNVFSKYADNHW--YLGIDIGYKQSMKKTNNHAFHHTQFSLTGKAFNL 1566
Qy      1265 GOLGIQPYGVNRYFIRENYSQSEVRAKVTSLAFNRVYNG--RVDTPTPTDNISVAPYF 1324
Db      1567 GNFGITPTVGRYSYLSNANFALADRIKNNPISYKTAAPQVLDSTY--HIGESVTPPL 1625
Qy      1325 FVNYVDVSNANVQTVNLTVLQOPFGRYQKEVGLKAEILHFIQSAISFSSQSGQKQ 1394
Db      1626 SARV--DTNQSQSKLNVQYDFAVYVENOQYNAGLKIKYHNVLSLIGLTKAQAEKOK 1684
Qy      1385 NNGVKIGY 1392
Db      1605 YAEKLSF 1692

RESULT 11
A41859
IGa-specific metalloendopeptidase (EC 3.4.24.13) type 1 precursor - Haemophilus influenzae
C:Species: Haemophilus influenzae
A:Variety: strain HK715
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 08-Dec-2000
C:Accession: A41859
R:Polisen, K.; Reinholdt, J.; Killian, M.
J. Bacteriol. 174, 2913-2921, 1992
A:Title: A comparative genetic study of serologically distinct Haemophilus influenzae ty.
A:Reference number: A41859; MUID:92234949; PMID:1373717
A:Accession: A41859
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-1702 <POU>
A:Cross-References: GB:M87489, NID:G143906, PIDN:AAA24966.1; PID:G148907
A:Experimental source: strain HK715
A:Note: sequence extracted from NCBI backbone (NCBI:P.97282)
C:Superfamily: IGA-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase

Query Match 22.1%; Score 1605; DB 2; Length 1702;
Best Local Similarity 26.4%; Pwd. No. 2.7e-70; Indels 550; Gaps 52;
Matches 479; Conservative 251; Mismatches 536;

```

```

Qy      6 FRILNFLLACISIGVSOAMAGHTYFGIDQYRDPFAENKGFYVAGQNIKYTNKQGLVY 65
Db      6 FKLNFALVAVAFALPTTEALVYRDVDYQIFRDPFAENKGRSFGATVVEVDKNNHSLG 65
Qy      66 TSMTKA--PMIDFSVVS--RNSVNAALVENQIVSVAH--NVGTYVDVFG--AEGN---NEDQH 117

```


Db 6 FKNFETALIVANALPYTEALVYRDVDYQIFRDPFAENKKSISVGAINTVEVDKKNOSG 65
 Qy 66 TSMTKA-PMIDSVVS-RNGVALVEXQIIVAH-NGVTTDVOFC-ABGN---NPDQH 117
 Db 66 SALPANGIPMIDSVVDVKRLATIVNPQIVGVKAVNSGSEIHFQNLKMNNGNNAKSH 125
 Qy 118 RFT-----YKIVKNNYKKNLHPY-----EDYNNPRLKZVTAPID-MT 159
 Db 126 RDUSSSENNRYTVEKNNPTEENVTSTTKEBODAQKREDEYMPRLKZVTAPIDAST 185
 Qy 160 SNNNGSTYSDRTKPEPERVITSGROF-----RNDQDQGDQ 195
 Db 186 ANNNKGETNSDKYPAFVRLGSSGQFIYKKSRYQLITTEKQQLNMDVGGDNLLEL 245
 Qy 196 VAGAYHYLTAQNTNOKAGNGSYLG-----GVYRAGEYGLPIAGSKGSSGNET 249
 Db 246 VGNAYTYGLAGPYKYNHNNGLIFGNSKEHSDPKGILSODPLNTYAVLADSSGSLFY 305
 Qy 250 YDAEKQKMLINGILREBNPEKENGQVLYRSYFDE--IFEDLITSLYTPAGNGYTTI 307
 Db 306 YDREKGMV-----FLGSYDFWAGNKKSNQEMNITKHEPAKTYQYISAGSLTG 355
 Qy 308 SGND-----NGGSTITQSGIPSHIKITLANSLPLKCKKVNHPYDOPNYSRLANG 362
 Db 356 SNTQYTWQATGISTITGGEPISVDLT-----DKD---KNNHG 392
 Qy 363 ETLYPMQKQSLIFASDINOAGAGLXFEENFTVSPNSQ--TWQAGIHVSNNSTYTWKY 421
 Db 393 KSITL--KSGGITLTMNHLDOAGAGLFFEBDEIVYKGTSDSTTWYAGAGSVAAKGTVMKY 450
 Qy 422 NGVEHDLKIGKGLHYQAKGNNKSGISVGGKYLEQOQDQGNKQAFSEIUNGSRG 481
 Db 451 HNPXYDLAKIGKGLHYBEGKNNKGLKNGGVTYILQKADANNKYNQAFSGYGVYSRS 510
 Qy 482 TVQANDKQPDITKFFGFRGGRLDNGSHLFFKXQNTBEGMNTNHTTQAANVTYTG 541
 Db 511 TVVANDKQVDPNSIYFGRGGRLDNGSHLFFKXQNTBEGMNTNHTTSHITTYG 570
 Qy 542 NESIVLN-----549
 Db 571 BSLLTNPVITTSYINBAODDHPRLRSIPYQVLYPNODNRSYTLKAGSTRSLPONS 630
 Qy 550 -----GNINLQDYRKELIANGFGETD--KXKNGKSLNLIYPTTED 590
 Db 631 GSNSEMLYMGRTSDAKKRYNMHNNERNANGNGYFGEERTKATQNGKLTNTFNGSKDQ 690
 Qy 591 RULISGNTLKDITQTKGLFFSSGPRPHAYNHLNKRNSBMGI-----PAGE 640
 Db 691 KRFLLTGSTLNGDLAVENGKTLFLSGRPHX-----RDLGISTYKXDPHTTENE 742
 Qy 641 IYMBHDNTNTPFAENFQKGSANVS-RNVSIBGNVTYSNNANAIFGVWXPQNTICT 699
 Db 743 VVEDDMLNKNFALTMNTGNASISYSGRVANITISITASNAQVHIG--KTDPLYCV 800
 Qy 700 RSDYGLTTCQKVDLDTQVINS:PKTQINGSLNLDNATNMGAKLNGVTLTNHSQ 759
 Db 801 RSDYGYTGENSNLS-KALNSNFPNLAGNNLTENS-----833
 Qy 760 FTLNVAOTIGNPIGNSSTATNTNANNGWHITNSAOFISKSHSHOIGSDKSTYVT 814

Qy 980 -----QELNDVVR-----AEQAEITLE-----397
 Db 1020 APSAQSNNEIARVETPPVPAPATASIASEQETPRALTAQAPAMEETNTANSITIAK 1079
 Qy 998 -----AKOVEP7--1004
 Db 1080 SPTATQENPNSSVSEETTERKVAENPPQENETVAANQEQAITEPTQNGVAKEQPTW 1139
 Qy 1005 AKTORGE-----1011
 Db 1140 ANTOQNETAQSEKTEBETQATKSEPTESVTSVENQPEKTVSOSTEDKVVKEKAKV 1199
 Qy 1012 -----PKYSRBARAA--FPDLPDQSLINALEAKQAEIATQOKAKATKVRSK 1061
 Db 1200 ETEETQKAPQVTSKEPPQAFEPAPPEVPTDT--NAEBAQALQOTPTVAABETSPNSK 1257
 Qy 1062 RA-----VFSD-----ELLQD-1072
 Db 1258 PAEBTQPESEKTNAPVPTVNSENTATQPTETETAKVEKENTQEVQVNASQSEFKQEP 1317
 Qy 1073 -----1072
 Db 1318 AAKPOAQTPQAPARENTVLTNANGEPQOPQOPQUSPAVEPTGEG-AANSKAPAPQAO 1377
 Qy 1073 -----SLPAA 1079
 Db 1378 AKQETPARENVSTWTKEPQOTSATVTEQPAKETSSNQBAPAPENSINTGSAETTYE 1437
 Qy 1080 ALEVITAPQO--SEKOR-----1094
 Db 1438 TAEKSDKQMETVTEENDQPEANTVADSVANNSESSSEKSRERSVSQPEKISAEITV 1497
 Qy 1095 -----LAQEBE-----KQKQ 1106
 Db 1498 ASQOETTVDSVSTPKERSRTRBQVQNSYTEPELEPTNNAEMENVOSSANNVANSQAL 1557
 Qy 1107 KDLISRTSNKLSLSA-----TVNSKLSVQDELDRFLV--DQAGSAVNTINADKRRYS 1160
 Db 1558 RNLTSKTNVAISNAKAKQFVALINQVAVSQHISQLENNEGQVNTVISTSNKXVSS 1617
 Qy 1161 DAFRAYQOQKTNLQOIGVQKALANG-RIGAVFSSHSQDTPBQVYKNEALTMMMSGFAQY 1219
 Db 1618 EGYTRPSSKSTQ-TOLQMDQTTSNVQUGVFTYVKSNNPD-KASSNTTLAQVNTYSKY 1675
 Qy 1220 -----QMGDLQGVNNGVGTISASMAEEOSEKIHKAIVGNASVQFPLQGLGQPYFCV 1275
 Db 1676 YADNEM--YIGILGIGYQSYLQTNNAKFAAHTAQIGTAGAPLGMFAKPYGV 1732
 Qy 1276 NRYFIERENYQSEEVRYKTPSLAFNRNAGIRVDYTTPTDNI SYKPYFNNYVDVSNAN 1335
 Db 1733 RYISYJNADPALADPRKINPISVKTAPQVUDLSYTY-HLGEFSTITPLSARY-DAPQGN 1790
 Qy 1336 VQITVNLVTLQOPFGRYWQEVGLKAEILHFQISAFSKSQSQGQXQNNGYVLGY 1392
 Db 1791 GKINVSVDYFAYINENQOQNAGLKIKYHNKVLISLIGLITXKQKQKTEVGLST 1847

QY 1027 -----TLPOSLINALEKO-----AELT----- 1045
 DB 1193 EKKAEELAKORAEAREBAQALAAARAKAEAEBAQALAEAREBAQALAEASQA 1252
 QY 1046 -----AETQ----- 1049
 DB 1253 AEAQALAAQOKALAQOEAREKAAELAVKOKETERKTAELAKORAAABAARQOEBAQ 1312
 QY 1050 -----KSKAKKKV-----BSKRAVFS 1066
 DB 1313 TAEALARQOEAREBAQALAEASQAQALAEAREBAQALAEAREBAQALAEAREBAQ 1372
 QY 1067 -----PLDPSLFALEAL----- 1081
 DB 1373 DYMAASQNRPRKRGKSTLPAPSPSPSSAAPPALANPDWYNDYEHPLDALEDEN 1432
 QY 1082 --EVIDAPQOSEKD-----RLAQEAE-----KOKKQKULISRYNSA 1117
 DB 1433 VSESVYDSDKQPODNTLHEKEVETSLQPRAPQPRQATOPQAOAAQDAVSTYNTSA 1492
 QY 1118 LSELATYNSMLSVODELD-----RLPVDQAQSAWNTNLAQOKRYVSDAPRAY 1166
 DB 1493 LSPAMASTQSTL-----LDTGASLSTRHIAQSKRADAEKUSVMSNIGYGRDYSAGQYKAF 1547
 QY 1167 QOQKTLRQIGVOKALA-NGRIGAVSHSRSDNTFDEQVYNHATLTPWAGPAQYOWADLO 1225
 DB 1548 SSKRTQ-TQIGIDRSLSENNQIGCVLTYSDSQHFD-QASGNQTVQANLTKGYLNDAM 1605
 QY 1226 F-GVNVGTGISHASMAEBSQRKIRKAINGVANASYQFRLQGLQIPYGVNRYFLEREN 1284
 DB 1606 YMAGDIGAGSLRSRLQOQKAMFNRTSICGTILGNTLKNQFELPVPAGIRYSRSSAD 1665
 QY 1285 YOSEHYRYKVTSLAENNYNAGIRVDYTFPPTDNISVQPY-----FVNVY-----VDV-SNN 1335
 DB 1666 YKJGNDVAVSVMSSVKTLTGGLDPAYRF-KYGNLTPBLSAIPYANQKGVNAGNSF 1724
 QY 1336 VQTTVNLTVLQOPPEGRWQKEVGLKAEILHFOISAFISKS-QQSOLQKQONWVKLGYR 1393
 DB 1725 VYRADN-----QQQYASGALLVRYNTLVNNGSITKQKQLEKQSGQIKQIR 1772
 QY 1394 W 1394
 DB 1773 F 1773
 RESULT 15
 154632
 Lsh protein - Escherichia coli
 C:Species: Escherichia coli
 C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 08-Dec-2000
 C:Accession: 154632
 R:Provenance, D.L., Curtiss, R.
 Infect. Immun. 62, 1369-1380, 1994
 A>Title: Isolation and characterization of a gene involved in hemagglutination by an avian
 A:Reference number: 154632; MUID:94178945; PMID:8132344
 A:Accession: 154632
 A>Status: preliminary; translated from GE/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1377 <RES>
 A:Cross-references: GB:U27423; NID:g469235; PTDN:AAA24698.1; PTD:g469236
 C:Superfamily: Iga-specific metalloendopeptidase
 Query Match 15.3%; Score 1113.5; DB 2; Length 1377;
 Best Local Similarity 26.2%; Pred. No. 1.9e-46;
 Matches 403; Conservative 197; Mismatches 519; Indels 421; Gaps 59;
 QY 3 KTVRLNLF-LTACISLIGIVSQAMAGHTFGIDYQYRDPAAENKGFVGAONIKVYKOG 61
 DB 29 KSVRLCFPVLLILIPVLSAGSLAGTVNNELGQFLRDPANRGMFRPATITAIYKOG 88
 QY 62 QLVTSMTAKPMIDSVV-SNGYALAVENQYTVVAVNNGYTDVDFGAEKGNPDQHRFT 120
 DB 89 EFVGT-LDKAMPDESADVSEIGVATLINFQYIASVKINGNTVNSFG-DGEN-----R 140

QY 121 YKIVKRNRYKONLHPYEDYENRBLHKFVTEAPIDMTS-NNNGSYSDRTKYPERVRI 179
 DB 141 YNIDRNNAF-----SLDFHAPRLDKLVEVAPTAFTAQAVAGAY-DKEEYVVFYRL 193
 QY 180 GSGKQFWRNDQKQVQAGAHVILTGNTNORAGNGYSLGGDVRKQ-----EY----- 231
 DB 194 GSGCYIKDSNQCQLQWKGASVWLTGGT-----VGSLSQYGEVISTSSGLVFDYKLN 247
 QY 222 GPLFIAGSGDSSPMFIYDAEKQKWLINGILRENGPEKENGFOLVKSYFDEL-F--E 289
 DB 248 GAMPFYEGAGDSGSLPAPFDIVQKKNVYGVLTANNGAGRGKNNMAYIILDLISQKREED 307
 QY 290 RDLHTSLYTAANGYTTLSGNDNQSSITQKSGIPSEIKITLAWMSLPLKEDKQVNHRY 349
 DB 308 NDAPVTFRTSEAGALWMSFNSSTGALITQ----- 337
 QY 350 DGEVNTSEFR-----LNGGETLYFMQKQSLIFASDINGQAGGVFEENFTVPSNQI 403
 DB 338 -GTTVYAMHQOQNDLMAKRLIFQOQ-NGQINLQSVSQAGSLFTRDNVYVTSNGST 395
 QY 404 WQAGIHYVSENSVYWKVNGYEHRLSKIGKTLHVAQKGENKSSISVGDGVILLEQAD 463
 DB 396 WTGAGIVVDNGSVYWKVNGYEHRLSKIGKTLHVAQKGENKSSISVGDGVILLEQAD 455
 QY 464 DQGNKQAFSEISGLVSGRGTVQINDKQFDTKPYEGFGRKLDLNGHSLTFKRIQNTDS 523
 DB 456 NKGQYQAFSSVNTASGRPTVVLTDEROVNDTVSWGKGTGLDNGSGSLTFKQLAGADYG 515
 QY 524 MATVNHNTQAANVTITGNESTVLPNANNKLDY-----RKEIAYNGM----- 567
 DB 516 AVLAN-VNDRATIT-----LDYALPDRALNNGSGSGKGTACNLK 557
 QY 568 -----FGE-DKXKH----- 576
 DB 558 YNRPYNTTQYETLXQSTGYFPPTDQSNATWSEVHGSGDAQVLVADRENTAGLTFHQ 617
 QY 577 -NGRLNLTLYK-PTTEDRTLLLSGGTNLKEDITQYKGLFSGRPTPYAHYK-----KR 629
 DB 618 LKGNLNVNRLPBGVGTALWMDGAAIDISGFTQENRKLTLQSHVHAYVYDSVADKLA 677
 QY 630 WSEMGATPQGEIYWDMDMINTREKAKENPOLKGSAAVNSASISIEGMVY-SNNANAFEG 688
 DB 678 SGDSHVLTQPTSPSQDEKMERSTFDRLSLANTDFGLGRNLT---NTTIQADNSVTLG 734
 QY 689 ---VVPNQ-QWTLICTRSDWTGLTTCQKYLDTDKVINSIPKQYNGSINT----- 734
 DB 735 DSRVFTIDNQGQGTAFITLEGTSVAT-----KADKSVENGTVNLNDSGVANI 782
 QY 735 -----TDNATANVKGIAKLNANTLTN-----HSQTLSENNAQTQGNIR 773
 DB 783 NDIENGCIQANNSTYVSSDAVLNGLSTLTALNLKNCANALASQSFVSDGPNISDA 842
 QY 774 LQDNS-----TATVDNANLN-----GNVHLTD----- 795
 DB 843 LSLNRPDEVSHTLLPVYDYAGSNLKKSDPRLANVPYMSLSGNINQDQGYTLTQGESE 902
 QY 796 -SAQSLAKNSHFHQLOGDKG-----TTLVLENAWTMPBCDTLONTLANSITF 844
 DB 903 LSPDLTLQNCQVLYSLRGYRANKSGSLNAPATVSTMDTQNMGNKSTFAGMMLNTIYV 962
 QY 845 LNSAVYASSNNTPRRSLETETTPSAEHRNTLTVNGKLSQGGTFQFISLGGYSDTL 904
 DB 963 FNGG-----TSP-----FTVLTDINDAVQSAFWRTDL--NKADTL 997
 QY 905 KLSNABG-DYILSVRYTGKPEPTELQGLT-----VESKQNPUS 943
 DB 998 VLNKSTVAGDSIYV-NFLKPSKQTLIDIPYASAPATADULPASRVVGSVDYPTL 1056
 QY 944 DKLFTLENDVDAKLYLVKNDGEFRLNPIKQELANDLVBAEQARTL-EAKQVE 1002
 DB 1057 SVAKEDGKGVAVDQ---YQVARNQGGKAAATFMMHSTNNFTITVNNLNKRMEDRLDIN 1113

```

Qy 1003 ZTAKTQGEPEKVSRRARAAPD-----TLPPDSILNALZAKQAEIETQ 1049
Db 1114 GEAGTWY--RLINGSGSADGFTDHTTILOMGADRKHELCSMDLFTGVMAITYTDTDSAD 1171
Qy 1050 KSKAXEKKQVSRKRAVSPDPLDOSLFALEMALEVTDAPOQSE-KDRIAQEAEKOROKD 1108
Db 1172 LYSKTKS-----WGGGYASGLFPRGAYFDVIAXYIHENKXYDLNPAAGAKONFR-- 1222
Qy 1109 LISRYNSALSLSATVNSMLSVDELDRLFYDQASAVTNTIAQDKRYDSDAFRAYQ 1168
Db 1223 -----SHSLYAGAEWGYRPHLT-----DTTFVEPOAEIWMGR-----QG 1257
Qy 1169 OKTNLRQIGVOKALAN-----GRIGAVFSHSRSDNTFDEOVKNHATLTMMSGPAQYCW 1221
Db 1258 QTFWANDSGMDVSMRNSVNPVYKRTGV-----SGKTFSGK--DMSLITARAQI-HYEF 1308
Qy 1222 GDLOFGVNVGTGISASKMAEEOQSKCHRKAINGVNASYOPRIG--OLGIOPYFGVNR 1278
Db 1309 -DLTDSADVHLKDPAQE-HQINRGKDSKMLYGVGLNA-----RFQDNVRLGLE----- 1354
Qy 1279 FIERENTQSEERAVKTPSLAFNR-----NAGIRVDYTF 1312
Db 1355 -VERS-----AFGKNTDIDAINNR--YSF 1377

```

Search completed: June 23, 2004, 17:32:38
 Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2004, 17:29:18 ; Search time 56 Seconds
(without alignment)
7854.150 Million cell updates/sec

Title: US-10-645-655-2

Perfect score: 7273
Sequence: 1 MKKVFRNLFLTACISLGLV.....SQGSQLGRQNVKLGIRW 1394

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 31518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5807	79.8	1436	2 08GCP3	08GCP3 haemophilus
2	5297.5	72.8	1391	2 08GCP2	08GCP2 haemophilus
3	4178	57.4	1449	2 09F325	09F325 neisseria m
4	4178	57.4	1449	16 09JMB4	09JMB4 neisseria m
5	4091.5	56.3	1444	2 07MILL	07MILL neisseria m
6	4085	56.2	1457	16 09JMB6	09JMB6 neisseria m
7	4084	56.2	1457	2 08GKS4	08GKS4 neisseria m
8	4072	55.9	1457	2 03X7H1	03X7H1 neisseria m
9	4067.5	55.9	1454	2 09F324	09F324 neisseria m
10	2523	34.7	507	2 08KRT9	08KRT9 haemophilus
11	2128	29.3	505	2 08KGS5	08KGS5 haemophilus
12	2087.5	28.7	510	2 08KSC0	08KSC0 haemophilus
13	1839.5	25.3	526	2 08KSB9	08KSB9 haemophilus
14	1724.5	23.7	514	2 08KND0	08KND0 haemophilus
15	1676	23.0	1431	16 09JMX3	09JMX3 neisseria m
16	1654.5	22.7	1561	2 051169	051169 neisseria m

17	1641	22.6	1552	2 09ALU6	09ALU6 neisseria m
18	1620	22.3	523	2 08KSC1	08KSC1 haemophilus
19	1620	22.3	523	2 08KIB7	08KIB7 haemophilus
20	1544.5	21.2	185	16 09KCB4	09KCB4 neisseria m
21	1538	21.1	1764	2 09JTB4	09JTB4 haemophilus
22	1536.5	21.1	1773	16 09JTB9	09JTB9 neisseria m
23	1432	19.7	997	2 09S6X2	09S6X2 neisseria m
24	1431	19.7	993	2 09S6X3	09S6X3 neisseria m
25	1421	19.5	997	2 030575	030575 neisseria m
26	1417.5	19.5	996	2 057309	057309 neisseria m
27	1412.5	19.4	996	2 030574	030574 neisseria m
28	1397.5	19.2	992	2 030573	030573 neisseria m
29	1397.5	19.2	992	2 09S6X4	09S6X4 neisseria m
30	1393.5	19.2	992	2 09S6X5	09S6X5 neisseria m
31	1193.5	19.2	992	2 057035	057035 neisseria m
32	1113.5	15.3	1377	2 047682	047682 escherichia
33	1109.5	15.3	1371	2 088093	088093 escherichia
34	1089	15.0	1371	16 08CWC7	08CWC7 escherichia
35	1085.5	14.9	1372	16 083079	083079 shigella fl
36	1083.5	14.9	1372	2 054151	054151 shigella fl
37	1083.5	14.9	1373	2 09AL58	09AL58 shigella fl
38	1060	14.6	1376	2 083W66	083W66 escherichia
39	1051	14.5	1376	16 08FRM0	08FRM0 escherichia
40	1023.5	14.1	1335	2 091A54	091A54 escherichia
41	1015.5	14.0	1335	2 091A58	091A58 escherichia
42	935	12.9	1364	2 08VSU2	08VSU2 shigella fl
43	935	12.9	1366	2 054165	054165 shigella fl
44	931	12.8	1364	2 099QC6	099QC6 shigella fl
45	924.5	12.7	1364	2 084GK0	084GK0 escherichia

ALIGNMENTS

RESULT 1

08GCP3 PRELIMINARY; PRT; 1436 AA.

AC 08GCP3: 31-MAR-2003 (PREMBLrel. 23, Created)

DT 31-MAR-2003 (PREMBLrel. 23, Last sequence update)

DT 31-OCT-2003 (PREMBLrel. 25, Last annotation update)

DE Adhesin (Fragment).

GN HAP.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Haemophilus.

CC

OX NCBI_TaxID=727;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PB60295;

RX MERLIN=22242173; PubMed=1235562;

RA Cutler D., Mason K.W., Howell A.P., Fink D.L., Green B.A.,
St Geme U.W. III;
"Immunization with Haemophilus influenzae Hap Adhesin Protects against
Nasopharyngeal Colonization in Experimental Mice".
J. Infect. Dis. 186:1115-1121(2002).

EMBL: AF66379; AAN7923.1; -
GO: GO:0004295; P-tyrosine activity; IEA.
GO: GO:0006508; P-proteolysis and peptidolysis; IEA.
InterPro: IPR006315; Autotransport.
InterPro: IPR005546; Autotransporter.
InterPro: IPR001254; Peptidase_S1.
InterPro: IPR000710; Peptidase_S6.
InterPro: IPR004899; Peptidase.
Pfam: PF03787; Autotransporter; 1.
Pfam: PF02395; IGH1; 1.
Pfam: PF03212; Peptidase; 1.
PRINTS: PR00921; IGASERPTASE.
TIGRFAMs: TIGR01414; autotrans_bar1; 1.
PROSITE: PS00135; TRYPSIN_SER; 1.
NON_TER 1436 1436
SEQUENCE 1436 AA: 159577 MW: CBB3950983BDB6CB CRC64:

Query Match 79.8%; Score 5807; DB 2; Length 1436;
 Best Local Similarity 79.3%; Pred. No. 2, 5e-242;
 Matches 1149; Conservative 78; Mismatches 154; Indels 68; Gaps 14;

QY 1 MKKTIVRLNFLTACISLIGVSOAMAGHTYFGIDYQYTRDAENKKGFTVGAQNTIKYKXQ 60
 DB 1 MKKTIVRLNFLTACISLIGVSOAMAGHTYFGIDYQYTRDAENKKGFTVGAQNTIKYKXQ 60
 QY 61 GQVYVSTKAKMIDFVSVBNGVYALVYNQYTVSVANNGYTDVDFGAGSNGDQHRFT 120
 DB 61 GQVYVSTKAKMIDFVSVBNGVYALVYNQYTVSVANNGYTDVDFGAGSNGDQHRFT 120
 QY 121 YKIVKNNYKXONLHPEEDYHNPLRHKFTVEAARIMTSMNGSTYSYDRTKYERKRG 180
 DB 121 YKIVKNNYKXONLHPEEDYHNPLRHKFTVEAARIMTSMNGSTYSYDRTKYERKRG 180
 QY 181 SGROFNRNDQ-----KGDVAGAYHYLTAGNTHNORGAGNGYSYLGDVAKAGEYGPAP 235
 DB 181 TGEQWMTDEBQKQSSKSMILADAYLMR IAGNTHSOGAGNGTVNLGDTKPKNNYGPAP 240
 QY 236 IAGSKDSSGPMFTYDAKOKMLINGI.MEGRPEKNGENFOLVRKSYF-DEIFERDLHT 294
 DB 241 TGVSFGSSSPMFTDALDKMLINGVLTQNGPESGANGPQLIRKMFYDNVFEDELPT 300
 QY 295 SLVTRAGNGVYITISGNDNGSS-ITQ--KSGIPSEIKITLANKSLPLKEDKVENPRY-- 349
 DB 301 TELPRBNKNGHFTSNNGGTGYTQNEKVSHPQKRYTQLEMLKEXK--EPYVAA 358
 QY 350 DGPNTVSPRLNNGEYLYMDQKQSSLI PASDINGAGGLYFESGNTVSPNSNQTQAGI 409
 DB 359 GGVVAYXPRLNNKNTYFGDRGTGTLTENNINQAGGLYFESGNTVSPNSNNAITMGAGV 418
 QY 410 NVSNSTVYTKNKGJEHRLSKTGKTLHQAQKSGNGSSVGDGVYILEQADQCNKQ 469
 DB 419 HVGSDSTVTKVNGVGHDRLSKIGKGLIHQAKSGNLSJSVGDGVYILEQADQCNKQ 478
 QY 470 AFSEIGVSGRGVQVQANDKQFDTDKFYFGFGGLDNGHSLTFKRIQNTDEGAMTVN 529
 DB 479 AFKVGIVSGRAVQVQANDKQFDTDKFYFGFGGLDNGHSLTFKRIQNTDEGAMTVN 538
 QY 530 NTFQANVTITGNESTVLP-NGNNINCLYRYEALVNGRGEETDKKNGRNLITYPTT 588
 DB 539 NTFQANVTITGNESTVLP-NGNNINCLYRYEALVNGRGEETDKKNGRNLITYPTT 598
 QY 588 EDRLTLSSGNTLKGDTQTKGKLFESGRPTPHAYNHLNKRKSMGEGIOGSIWMDH 648
 DB 599 EDRLTLSSGNTLKGDTQTKGKLFESGRPTPHAYNHLNKRKSMGEGIOGSIWMDH 656
 QY 649 NRTKAFNFOIKGSAVVSRSVSSIEGNTVSNNAANATFGVVPQONTICTRSMTGLTT 708
 DB 657 TRTKAFNFOIKGSAVVSRSVSSIEGNTVSNNAANATFGVVPQONTICTRSMTGLTT 716
 QY 708 CQKVDLDIVYINSIPKQVINGSINTLQNAITANVGLAKLNGVYLTNHSQTTLSNNA 768
 DB 717 CKYVDLDIVYINSIPKQVINGSINTLQNAITANVGLAKLNGVYLTNHSQTTLSNNA 776
 QY 769 IGNRLSDNSTATVDNANLNGVYLTDSAFQSKSHSHQIOGDKGTVTEAATWMP 828
 DB 777 TGNLQSNHNAATVDNANLNGVYLTDSAFQSKSHSHQIOGDKGTVTEAATWMP 836
 QY 828 SOTTLQNLTLNNTITLNSAVSASSNNTTP-RRSLTETPTTSAHRFNTLTVNKLSCQ 887
 DB 837 SOTTLQNLTLNNTITLNSAVSASSNNTTP-RRSLTETPTTSAHRFNTLTVNKLSCQ 896
 QY 888 GTFQFTSSLFQYSDKLSNDAGCYTILSVRTGKEPTLEQULTVSKDQNPISDKX 947
 DB 897 GTFQFTSSLFQYSDKLSNDAGCYTILSVRTGKEPTLEQULTVSKDQNPISDKX 956
 QY 948 FTLEKDVADGALRYLVNDQFRLNRPKEOELENDIVYABOELETLAEOVEPTAKT 1007
 DB 957 FTLEKDVADGALRYLVNDQFRLNRPKEOELENDIVYABOELETLAEOVEPTAKT 1016
 QY 1008 QTGEKTVSRBRARAFAFDTLTLDQSLNLALEKQAELETAETQSKATKTVKSKRAVSD 1067

DB 1017 QVSNARVRSK---RAVSDTLTLDQSLDQVLAQVAPFAEKQKKA--KRVSKRAVSD 1071
 QY 1068 PVLDS-----LFALEALVETD 1085
 DB 1072 TLTPDQSLDQVLAQVAPFAEKQKKA---KRVSKRAVSD 1131
 QY 1086 APOQSEKRLALAEAKRKOKDLISHSNSALSELSATVNSMLSYQDELDRLFVDQMS 1145
 DB 1132 APOQSEKRLALAEAKRKOKDLISHSNSALSELSATVNSMLSYQDELDRLFVDQMS 1188
 QY 1146 AVNTMIADKRYSDAFRAVYQOQKTLRQIVQKALANGIIGAVFSHSKSDTDEQVK 1205
 DB 1189 AVNTMIADKRYSDAFRAVYQOQKTLRQIVQKALANGIIGAVFSHSKSDTDEQVK 1247
 QY 1206 NHATITMMSGFACQWGDLOFGVNVGTGISASKMAEGRXIRKALNVGNASVQFRLG 1307
 DB 1248 NHATITMMSGFACQWGDLOFGVNVGTGISASKMAEGRXIRKALNVGNASVQFRLG 1307
 QY 1266 QLGIOFYGVRVPIRENYQSEEVYKTPSLAFNRVYAGIRVDTPTDNI SVKPYFF 1325
 DB 1308 QLGIOFYGVRVPIRENYQSEEVYKTPSLAFNRVYAGIRVDTPTDNI SVKPYFF 1367
 QY 1326 VNVYVSNANVQTVNLTVLQOPFGRYQKQVGLKAEILHFQISAFISKSQSQSLQKQON 1385
 DB 1368 VNVYVSNANVQTVNLTVLQOPFGRYQKQVGLKAEILHFQISAFISKSQSQSLQKQON 1427
 QY 1386 VGVKLGYN 1394
 DB 1428 VGVKLGYN 1436

RESULT 2

Q86Q2 ID Q86Q2 PRELIMINARY; PRT; 1391 AA.
 AC Q86Q2; 01-MAR-2003 (TREMELREL. 23, Created)
 DT 01-MAR-2003 (TREMELREL. 23, Last sequence update)
 DT 01-OCT-2003 (TREMELREL. 25, Last annotation update)
 DE Adhesin (Fragment).
 GN HAP.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM K.A.
 RC STRAIN=TN106;
 RX MEDLINE=2242173; PubMed=12355362;
 RA Cutter D., Mason K.W., Howell A.P., Fink D.L., Green B.A.,
 St Gene J.W. III;
 RT "Immunization with Haemophilus influenzae Hap Adhesin Protects against
 RT Nasopharyngeal Colonization in Experimental Mice."
 RL J. Infect. Dis. 186:1115-1121(2002).
 DR EMBL; AF69380; AAN37924.1; -
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR006315; Autoctransport.
 DR InterPro; IPR005546; Autoctransporter.
 DR InterPro; IPR009203; Cys_Ser_trypsin.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR004899; Peptidase_S6.
 DR Pfam; PF02397; Autoctransporter; 1.
 DR Pfam; PF02395; IGA1; 1.
 DR Pfam; PF03212; Pertactin; 1.
 DR PRINTS; PR00921; IGASERPTASE.
 DR TIGRFAMs; TIGR01414; autoctrans_bac1; 1.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 FT NON TER 1391 1391
 SQ SEQUENCE 1391 AA; A55385D8BC156FD1 CRC64;

Query Match

72.8%; Score 5297.5; DB 2; Length 1391;

Best Local Similarity 74.7%; Pred. No. 2e-220;
Matches 1052; Conservative 109; Mismatches 216; Indels 31; Gaps 15;

```

QY 1 MKKTVFRLNLTACISLISVSOAMAGHTYGIYQYRPANRKGKTTVGAOKIKYKQ 60
DB 1 MKKTVFRLNLTACISLISVSOAMAGHTYGIYQYRPANRKGKTTVGAOKIKYKQ 60
QY 61 GOLVGTSMTKAMIDSVSRNGVAIVNOYIVSAHVNGYTVDEGAEANNPDQHRT 120
DB 61 GEMIGTMMKGVMPDLSMWVRGYSYTLISGHLISVAHVGVDOVDGEMGENPDQHRK 120
QY 121 YKIVRNNYKKNLHPEYEDYHNPKLHKTETAPLIDMTSNMGSTSDTKPEPVKIG 180
DB 121 YKIVRNNYKKNLHPEYEDYHNPKLHKTETAPLIDMTSNMGSTSDTKPEPVKIG 180
QY 181 SGRTMRNDQK--GDVGAAYHVLTAQTNRNGAGNAYSLGSDGVRKAGEPLITG 238
DB 179 SGHMMKDKNNKTIGDIAYG--SMILGNTFEDGPNNGTLELNGRVQMPKGTPLTNG 237
QY 239 SKDGSQPMFTYDAEKQWLLINGILREGNPEPEKENGQOLVRSYDEIEERDLHTSLYT 298
DB 238 SPEDSSPMFTYDKEVKKWLLNGVLRBGNPYAAVNGSYQITKDYFOGILNDITANFD 297
QY 299 RAGNGVYITSGNDG-----QSTQKSGIFSEIKITLANMGL-PLKCKDVANPEYDGP 352
DB 298 TNAEYKFNIGSDHNGRVATIKSTLPKKAIOBERIVSLYDINSOLHDPARDKSDSPYKGP 357
QY 353 NIYSPLANNGETLYFMOKQSLIFASDINGAGGCLYFSGNFTVSPNSQ--TWOGAGIHV 411
DB 358 NPWSPLMHGKSIYFGDGTGTLTENNINAGAGGLTFBGNRYVGNKNNITWQAGVAV 417
QY 412 SENSTYWKVNGVENHRLSKIGKGLYHQAENGKGSIVGKXVILEQADQGNKQAF 471
DB 418 GESSTYQVYHNPBGDLSTKIGLTLVNGKGNLGLSLGKGLVLLDQADSSGQKQAF 477
QY 472 SRIGLVSGRGTVQIANDKQPTDKRYFGFRGGGILDLNGSLTFKAIQNTDEGAMTYNAT 531
DB 478 KEVGIYSGRATVQIANSADQYDVPNNIYFGFRGGGILDLNGSLTFKAIQNTDEGAMTYNHA 537
QY 532 TOAANTTNGESIVLNGNNINKLDYKEITANVGMGENDNKNNGNLNIYKPTEDR 591
DB 538 SOTANTTTGNATL---NSDSKQNLNKKDLAFNMGFBGDDAKNNGNLNNOPTNMAEN 593
QY 592 TLLSGGTLKGDITOTKGLFPSPGPTPHAVNHLNKRWSBEGIPQBIYWDHDMINT 651
DB 594 HLLSGGTLNGLNITLQNGGLTVPSGRPTPHAVNHLNKRWSBEGIPQBIYWDHDMINT 653
QY 652 PKAENFOIKGSAYVSRVSSIEGKWTYSNNANATPGVYNQANTICTRSWGTCLTTCK 711
DB 654 PKAENFOIKGSAYVSRVSSIEGKWTYSNNANATPGVYNQANTICTRSWGTCLTTCK 713
QY 712 VDLTKKVTNLSIPKTOINGSLNLTDAATANKGLKNGVNTLNHSGFTLSNNAVOIG 771
DB 714 VDLTKKVTNLSIPKTOINGSLNLTDAATANKGLKNGVNTLNHSGFTLSNNAVOIG 773
QY 772 IRLSDNSTAYDNANLNGVNLTDLSAQSFLNNSHSHQIGDKGTTVLNENATVWPSDT 831
DB 774 IRLSDNSTAYDNANLNGVNLTDLSAQSFLNNSHSHQIGDKGTTVLNENATVWPSDT 833
QY 832 TLQNLNNSSTTLNLSAYSSANNLT--RRSLSETTTTPSAHRPNTLTNGCLSGQGT 889
DB 834 TLQNLNNSSTTLNLSAYSSANNLT--RRSLSETTTTPSAHRPNTLTNGCLSGQGT 893
QY 890 FQFTSLFGYKSDKLTLSNAEDGYLISVNTGKEPPTLSQTLVYESKDNQPLSKLKEF 949
DB 894 FQFTSLFGYKSDKLTLSNAEDGYLISVNTGKEPPTLSQTLVYESKDNQPLSKLKEF 953
QY 950 LENDVDAGALRYKLVXNDGEPFLNPIKEGHLNDIVARAEQERTLAKQVEPTACTQT 1009
DB 954 LENDVDAGALRYKLVXNDGEPFLNPIKEGHLNDIVARAEQERTLAKQVEPTACTQT 1013
QY 1010 GERKRSRAAANAAPPTLDQSLNALAEKQAEITLFTQSKYAKTKYVRSKAV--FSQ 1067
DB 1010 GERKRSRAAANAAPPTLDQSLNALAEKQAEITLFTQSKYAKTKYVRSKAV--FSQ 1067

```

```

DB 1014 SKARVSR--RAVESDPLPQSLKALPAKQ-LITETQTSKA--KKRYSKARAREFSD 1267
QY 1068 PLTDSIFALEALAEVIDAPQOSEKRLAQBAAEK-OROKDLISRYNSALSELATVN 1126
DB 1068 TLTPQ--LIGALAEVIDAQQVKEPQTOEEERQORKEELISRYNSALSELATVN 1124
QY 1127 SMUSVQDELDRLFTDQASAVNTNIAQDKRYSDAFPAVQOQKTLRLQIQYKRLANG 1186
DB 1125 SMUSVQDELDRLFTDQASAVNTNIAQDKRYSDAFPAVQOQKTLRLQIQYKRLANG 1183
QY 1187 IGAVFSHSRNDTPDQVQNHATLMMSGFACQWGDLOGVAVNGTGISASGAABEQSRK 1246
DB 1184 IGAVFSHSRNDTPDQVQNHATLMMSGFACQWGDLOGVAVNGTGISASGAABEQSRK 1243
QY 1247 IHRKAIYGVNAYQFRLGQIGIOPYGVNRYFIERENYQSEBEVYKTESLAFNRNNGI 1306
DB 1244 IHRKAIYGVNAYQFRLGQIGIOPYGVNRYFIERENYQSEBEVYKTESLAFNRNNGI 1303
QY 1307 RVDYTFPTNTNISVPRYFFNVYDVSNANVQTTNVLTVLQQPGCRITWQXVGLKAEILHF 1366
DB 1304 RVDYTFPTNTNISVPRYFFNVYDVSNANVQTTNVLTVLQQPGCRITWQXVGLKAEILHF 1363
QY 1367 QISAFISKQSQSLQKQONVGRKGYRW 1394
DB 1364 QISAFISKQSQSLQKQONVGRKGYRW 1391

```

RESULT 3

09F325 PRELIMINARY; PRT: 1449 AA.

AC 09F325, 01-MAR-2001 (TREMREL, 16, Created)

DT 01-MAR-2001 (TREMREL, 16, Last sequence update)

DT 01-OCT-2003 (TREMREL, 25, Last annotation update)

DE Adhesion penetration protein.

GN App.

OS Neisseria meningitidis.

OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;

OC Neisseriaceae; Neisseria.

OX NCBI_TaxID=487;

RY [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Z3515;

RA Ala-Aldeen D.A.;

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Z3515;

RA Abdel Hadi H., Wooldridge K.G., Robinson K., Ala Aldeen D.A.;

RT "Identification and Characterisation of App: an Immunogenic Autotransporter Protein of Neisseria meningitidis.";

RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ296276; CAC14670.1; -

DR MEMOS: S06.006; -

DR GO: GO:0008233; F:peptidase activity; IEA.

DR GO: GO:0004295; F:trypsin activity; IEA.

DR GO: GO:0005508; P:proteolysis and peptidolysis; IEA.

DR InterPro: IPR006315; Autolysinport.

DR InterPro: IPR005546; Autolysinport.

DR InterPro: IPR009603; Cys Ser trypsin.

DR InterPro: IPR001254; Peptidase_S1.

DR InterPro: IPR00710; Peptidase_S6.

DR InterPro: IPR004899; Peptidase.

DR Pfam: PF02395; IGA1, 1.

DR Pfam: PF02395; IGA1, 1.

DR Pfam: PF02312; Peptidase; 1.

DR PRINTS: PR00921; IGASERPTASE.

DR SMART: SM00020; TRP_SPC; 1.

DR TIGRFAMs: TIGR01414; autotrans_bart; 1.

DR PROSITE: PS00135; TRYPSIN_SER; 1.

KW Hydroxylase; Protease; Serine protease.

SEQUENCE 1449 AA; 159071 MW; C07850BD4347C25D CRC64;

Query Match	57.4%;	Score 4178;	DB 2;	Length 1449;
Best Local Similarity	56.2%;	Pred. No. 4.2e-172;		
Matches 816;	Conservative 214;	Mismatches 325;	Indels 98;	Gaps 17;

Qy	10	FLRACISGIVSMAACHYTPGIDYQYRPPAANGKFTVAKMIVKYNQOGLVGSMT	69
Dh	27	YLAICSPFELPQMAQHITPGINQYRPPAANGKRAVAKHOLEVKKKELVGSMT	86
Qy	70	KAMIDFSVSRNVAALYENQYIVSVANVGYTVDFGABGNPDQHSFYTKIVKNY	129
Dh	87	KAMIDFSVSRNVAALVGDQYIVSVANVGYTVDFGABGNPDQHSFYTKIVKNY	146
Qy	130	KONLHPEDDYENVRILKFTVBAAPIDHSMNGSTISDITKIPERVATSGGQFRRND	189
Dh	147	KPNHSHPYMGYHMERLHKFTVDAPPEMSTDMRGATISDEXKPERVATSGHHYRYD	206
Qy	190	QDGGDVGAHYLTANTHNRQAGNGYSYLGDVYRAGYGPFLIAGSKDSDGSPMF	249
Dh	207	DDHGDPLTISGAMLIGNTHMQGMNGNVSLSGDVREANIDYGPMLIAGADSDGSPMF	266
Qy	250	YDAEKOMLINGSILRBGNPFEGKENGQOLVRKSYR-DEIFERDHTSLYTRAGNYTIS	308
Dh	267	YDKNNKMLINGVLTQYPSYSGRENGFOLIRKOMEFDIDYGDHTVTFEPESNGHSEF	326
Qy	309	GNNQGGSTQKSGIPEIKITLAMS---PKEKEDVHNPR--DSGNYSPLNKG	363
Dh	327	SNNGGVTLENKATSNPLVOTVRLEPESINETDK--EPVYAGVNOQRRPLNKG	384
Qy	364	TLVEMOKGSLIASPDINGAGGLYFEGNFVSPNSQNTQOAGLHVSNGTYTWKVN	423
Dh	385	NISFIDYNGKGLISNNINGAGGLYFEGNFVSPNSQNTQOAGLHVSNGTYTWKVN	444
Qy	424	YEHDLSTKIGGTLVQANGKSGISYEGKTYLEEQADQGNKQAFSEILVSGRTY	483
Dh	445	VANDRLSTKIGGTLVQANGKSGISYEGKTYLEEQADQGNKQAFSEILVSGRTY	504
Qy	484	QUNDEKQPDTPKFPESRGGRLDNGHSLYFLKQNTBEGMIVNNTQOANTIGN	543
Dh	505	QUNDEKQPDTPKFPESRGGRLDNGHSLYFLKQNTBEGMIVNNTQOANTIGN	564
Qy	544	SIYLPNGNNIKMDLYRKEIAYNGMGETDKNKNRNLMIKPYTBDRITLLSGGTLNG	603
Dh	565	SIYLPNGNNIKMRNYSKEIAYNGMGETDKNKNRNLMIKPYTBDRITLLSGGTLNG	624
Qy	604	DIYQYKGLFPSSGPTPHAYNHLANKRMBEBSGPOGETYWHQDMINTKAKANQYQGS	663
Dh	625	NITQYKGLFPSSGPTPHAYNHLANKRMBEBSGPOGETYWHQDMINTKAKANQYQGS	684
Qy	664	AVYSRWVSIIEBNTVSNNAANATPGVWVNOONTICRSPWMLGTQCKVDLIDTRKYSI	723
Dh	685	AVYSRWVATVEBDHLSNNAQVFGVAFHQSHTICRSDWGLTINCVEKTIITDKVYAS	744
Qy	724	PKTQINGSINTLTDNATNAYKGLAKINGVYTLNHSOFTLSNNAITQGN-----	771
Dh	745	TKTIDISGNVSLADIAHNLITGLATINGLSANDTRYVSHNATQNGNLSLVGNAQATEN	804
Qy	772	-----IRISDUSTAYDNANINGNHLTSDAQFSLKS	804
Dh	805	QATLNGTNSAGNBSFNLNNAQNGSLTSDAKNANVSHSALINGVSLADQAFHENS	864
Qy	805	HFSHOLQDQKGTVLTENATMTMPSDITLQNTLNNSTYTLNSAV-----SASGNT	866
Dh	865	RFGQGLSGSKDTPALHLDGEMTLLPGTEIGNULMDNATITLNSAYRHDAAQNGQSVSD	924
Qy	867	PR---RRSLFETTPYSAEKRNITLTNGKLSQGGTFQPTSSLGYKSNKLTLSNDAED	913
Dh	925	PRRRSRRLSLSTPPTSVESRNTLTNGKLNQGGTFRFSYSELGYSRDKLTLESSBGT	984
Qy	914	YIISVANTQKEZETJBLQTLVSSXONQPLSDXKFTLENDVADGALRKYLYONDGEFRL	973
Dh	985	YTLAVANTQNEPYSLDLTVYEGKONKLSBNINPTLQNEHVHDGAMRQQLIRRDGEFRL	1044
Qy	974	HNPIKEQELNDLVRAE---QAEK-----TLEAKQYBPAKTQY-GEFYKRSFRAARA	1022

```

Db      1045 HNNVYGEDELSTGKAKHAKQAEKDNAGSLDLIAARDAALENTESVAEP-----ARQ 1097
Oy      1023 AFDPTLLPDGSILNALLDAKAQALTEFVQSKATKTAVSRKAFVSODPLLDQSFLPAALIE 1082
Db      1098 AGGEENV--GIHQDEEEK-RVQADKQSMALAQRAEABTRPATAAP-----PARA 1144
Oy      1083 VIDAPEGSEKNRLAOJEBAEK-RKQCHLISRYNSALSSELATVMSLSYODELDLPVD 1141
Db      1145 RRLPE-----POQPOPQPQRQLSRANSGLSEFSATINSVFANODELDVPFE 1197
Oy      1142 QAQSAVTNLINADKKREYDSDFRAYDQDKTLNRIGVQAKLANGRTGAVSHSRSDNTED 1201
Db      1198 DRNNAMWTSIGIRDTKHYSRQDPFRAYRQQ-TDLRQCIMQKMLGSGRGVILLFSHRNTENTED 1256
Oy      1202 EGVKHAHTLTLMSSPEFQKYQMEDLFQNGNVTGISAKSMAEQSEKTHKHALNYGWASQ 1261
Db      1257 DGGNSARHLAAGAFAFOYIGIKREFDIGISTGAFFSSGSLDGIGZKKIRRVLHGIDQARY 1316
Oy      1262 FRIGQLGIQPFQGNVNFPIERENYQSEEEVNPTELAFNNYNAGIRVLYTFPTDNIYSK 1321
Db      1317 AGCGGAGIEPIETGATNYFPGQADRYENNVIATPELAFNMYRAGIKADYSKPAQHISIT 1376
Oy      1332 PYFFANYVDVANAVQTNTLTLYIQPFGRYQKVGVGKKEIHLIQIALSFMSQSSQ 1381
Db      1377 FYTSLSTYDAASGKVRTRVNAVLAQDPFKTRSAAEWNAELKGLTSLHAARKGPQLE 1436
Oy      1382 KQONNVXKLGTYW 1394
Db      1437 AQHSAGILGTGW 1449

RESULT 4
OJUMB4 PRELIMINARY; PRT, 1449 AA.
ID OJUMB4-
AC OJUMB4-
DT 01-OCT-2000 (TEMBLrel, 15, Created)
DT 01-OCT-2000 (TEMBLrel, 15, Last sequence update)
DT 01-OCT-2003 (TEMBLrel, 25, Last annotation update)
DE Iga-specific serine endopeptidase [EC 3.4.21.72].
GN NS160457.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
CC Neisseriaceae; Neisseria.
CX NCBI_TaxID=65699;
RN RN
RP RP
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RM MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies L.M., Davis P., Devlin K., Felwell T., Hamlin K., Holtroyd S.,
RA Jagels K., Leather S., Moule S.K., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds W., Skelton J.,
RA Whitehead S., Spratt B.G., Barrett B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506(2000).
RU EMBL, AL162753; CB883754.1; -.
DR PIR, B81963; B81963.
DR MEROPS, S06_006; -.
DR GO, GO:0016787; F:hydrolase activity; IEA.
DR GO, GO:0004295; F:trypsin activity; IEA.
DR GO, GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro, IPRO06315; Antictransport.
DR InterPro, IPRO05546; Autotransporter.
DR InterPro, IPRO09003; Cys_ser_trypsin.
DR InterPro, IPRO01254; Peptidase_S1.
DR InterPro, IPRO00710; Peptidase_S6.
DR InterPro, IPRO04893; Peractin-.
DR Pfam, PF03797; Autotransporter; 1.
DR Pfam, PF03395; IGAL1; 1.
DR Pfam, PF03212; Pertactin; 1.
```

DR PRINTS; P000921; IGASERF.PASB.
DR SMART; SM00020; TYPE_SRC; 1.
DR TIGEPFAMS; TIGER01414; autotransp_bar1; 1
DR PROSITE; P000135; TRYPSIN_SEQ; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 1449 AA; 159070 MW; B858D0

Query Match	57.4%;	Score 4178;	DB 16;	Length 1449;
Best Local Similarity	56.2%;	Pred. No. 4.2e-172;		
Matches 816;	Conservative 214;	Mismatches 325;	Indels 98;	Gaps 17;

```

0Y 10 LTAACSLAIVSAQMAAHYFEGIDYQYRPAPAKCKEYVQAKYKYNKQGLVGSMT 69
Db 27 YLAI CSFGLIEQAPAHGHTFGELTYQYRDAKKNKRAVGAQDIEYVYKKSGLVGSMT 86
0Y 70 KAPMIDFSVSNHGAALVENQIVTSVAMHGYTVDVGAQGNPNQDHFYTKVKNXY 129
Db 87 KAPMIDFSVSNHGAALVDDQYIVSAHNGYNNVDFGAQRPNQDHFYSQYVKNXY 146
0Y 130 KKNLHPEYDDYHNHPRJAKVTYEAAPLIMTSNAGSTSDTKPERYRISGQCFHRND 189
Db 147 KEDNSHPHYGDIYHNERLHAKVTYDAEPVEMSDMRGNTYSDXEKPERVHISGHHYKXD 206
0Y 190 QDKCDOYAGAVHYTLAGNTHNQAGNCSYSLAGDYRAKSGYGLPLAGSKDSSPMFJ 249
Db 207 DDEHGGLSVSGAMLLGENTHMQMGNNVYVLSGDYRANXNDYGMPLAGAADSSPMFJ 266
0Y 250 YDAEKQKMLINGILREBNPEPEKENGFOYVKSYP-DEIFERDIHLSLYTRANGVYTTLS 308
Db 267 YDKTNKMLINLVQTYGYPYSRGENGFOYLRKQWPFYDDIYGDGHHYTFPEPSMAHSPF 326
0Y 309 GNDQGGOSITQKSGSIBELKITLANKSL--PLXEDKXHNHPR--DGNITSRLNNGS 363
Db 327 SNNGSGYVTEYTERNEKVNPLKVTYRFPDSILNETDK-BPVYAAQVGNQRPPLNNGS 384
0Y 364 TLYPMQXQGSILFASDINGAGGLYEFENETVSPNSNQVQAGLHVSENSTVYMKNG 423
Db 365 NLSTIDYNGKULLSNININGAGGLYFEGDPTVSPENNETHQAGVHISEDSTVYMKNG 444
0Y 424 VEHDRLSKIGKCTHVOAKENKGSISYGDGVILEQQADQNGKQAFSEILVSGYGLV 483
Db 445 VANDRLSKIGKCTHVOAKENKGSISYGDGVITYLDDQADQNGKQAFSEILVSGYGLV 504
0Y 484 QUNDQOFTDVEYFEGRGRDLZNGHSLFPRKQNTDEGAMTYNHTYQAAVYITGN 543
Db 505 QUNDQOFTDVEYFEGRGRDLZNGHSLFPRKQNTDEGAMTYNHTYQAAVYITGN 564
0Y 544 SYLVENGNINLQYRKEILVYNGMFGETDKKNGKGLNIYKPTEDRTLLSGGYNLKG 603
Db 565 SYLVENGNINLQYRKEILVYNGMFGESQDTYKNGKGLNIYQPAEDRTLLSGGYNLKG 624
0Y 604 DITDTGKLFPESGRTPHAYNHLAKRMBEMEGIPOEELYMHDIMYTRKANPQIGSGS 663
Db 625 NITYTGKLFPESGRTPHAYNHLGSMKMEGIPQELIYMNDDIMINTFKANPEIDQSG 684
0Y 664 AVYSRNVSSIEBNMTVSNMNAATPCVYMNQONTICRSDPWGLTTCQKYLDTYKVINSI 723
Db 685 AVISRNVAVBEDMHSNHAQVFGVAPHQSHIICRSMWGLNCEKTIIDDXVIALSL 744
0Y 724 PKTOINGSINLTNNATNVKGLAKLNGVTLNHSOFTLSNNATQCN----- 771
Db 745 TKTDISGNVSLADHAILNLTGLATLNGMLSNAGCTRYTVSHRAQONQNLSTVGNQATFN 804
0Y 772 -----TRSDNSTVYNNALINGNHLTDSAQSLKNS 804
Db 805 QATLNGTNSASGNSMFLSNNAQNGSLTSDNAKNAVSHSLNGNVLSDAAYVHEHNS 864
0Y 865 HFHQIQXQKGTYYLLENATMTWPSYTLQWLLINNSYTLTANSAY-----SASSNNT 856
Db 865 RFTQQLSGSDYTLALXDSQEWLPSSTELANGLNDAVITLLNSAVRHDAAGQTSYSDT 924
0Y 857 PR--RRSLFTTYSAEHRYTLTVNGKLSGGGTFOPTSLFGYSKDKLSNDBED 913

```

```

Dh 925 PBRBRBRSLBLSVTPPVSVSBRFNTLTVNGKNGQOTPREFSELFGYRSDKLIKLAESSEGT 984

Qy 914 YLISVRNTGKSEBTELQOLTIVESYDNOPLSDCKLTKTLENDHVDAGALVYKLVKNDSEFPL 973
Dh 985 YTLAVNNTGNEPVSJDLQTLVVEGDKNPLSENLPTLONHEVDAGAMVEYQOLIRKDEBEL 1044

Qy 974 HNTKEOGLHNDLVRAE---QABE-----TLAKNGVEPLAKTOT-GEPLVSRBAARA 1022
Dh 1045 HNVKVEQELSDKJGKBEAKQERKDNQSLDMLIAGRLDAEKTESVAEP-----ABQ 1397

Qy 1023 APFDTLPOSLNALBAKQZELTLETQTSKAKTKKVSNAKAVESPJLDQSLFALBALE 1082
Dh 1098 AGEENV---GIMQZBEERK-RVQJLDSDSLAKQGRATERRATTAPF-----RABRA 1144

Qy 1083 VIDAPQOSEKDRLAQBEAEKQ-RQKQULISRYSNSALSELATONSMVSYODEIDLFPVD 1141
Dh 1145 RROLPLQ-----PQVQPQPQOPQBRDLISRYANGSSEFSATLNSVAVQDEIDRVFAE 1197

Qy 1142 QPASAVNTNLIODKRRYRSDAFRAVQOQKTNLROIGYQVKKLANGRIQAVFHSRHSNTPD 1201
Dh 1198 DRRNAVMTSGIRDTKRYRQOPFRAYRQO--TDLROIQGQXKLGSGRVGLFESHKRENTYD 1266

Qy 1202 EOVKNKATLTVMGSGPQYQWQMDLQFGVNTQGTISASMAEESRKXHKERKAVNGSVASQ 1261
Dh 1257 DQGSNARLAAGAVFQGYQIGRFDIGISTGAGFSGSGSLSDGIGKIRRPVLYHYIQARIR 1316

Qy 1262 FRLGQZGLOPYGVNRYFIRENTYQSEBEVRVKTPSLAFNRKNAGIRVVDYTFETPTNISVK 1321
Dh 1317 AGSGGAGIEPIGATNRYFVQKADRYENVNVIATPELAFNRIRRGIKADYSEKPAQHSITC 1376

Qy 1322 PFFVNVYVDVSNANVOITVNLTLTQOPFGHVMQKVGILKAEIILHPOLSAFISKSQSGS 1381
Dh 1377 PYLSLTYDDASGQVTRVNTVLAADOPGKTRSAEMGVNAEIKGFTLSLHRAAKKGPQLE 1436

Qy 1382 KQOMNGVTKGYRW 1394
Dh 1437 AQBSAGIKKGYRW 1449

```

```

RESULT 5
OTM011
ID OTM011 PRELIMINARY; PRT; 1444 AA.
AC OTM011;
DT 01-OCT-2003 (TREMBLrel_25, Created)
DT 01-OCT-2003 (TREMBLrel_25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel_25, Last annotation update)
DE Adhesion and penetration protein (Fragment) .
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2996;
RA Serruto D., Adc-Bodie J., Scarselli M., Veggi D., Piza M.,
RT Neisseria meningitidis App, a new adhesin with autocatalytic serine
RT protease activity";
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=2996;
RA Bambini S., Rappapoli R., Piza M.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF56265; AAC08950.1; -.
FT KON TER 1
SQ SEQUENCE 1444 AA; 158147 MW; 33A147B91EE118EB CRC64;
Query Match 56.3%; Score 4091.5; DB 2; Length 1444;
Best Local Similarity 55.6%; Pred. No. 2,3e-166;
Matches 809; Conservative 213; Mismatches 335; Indels 99; Gaps 19;

```


PRINTS: PR00921; IGASERPTASE.
DR TIGREPMs; TIGR01414; autotrans bar1; 1.
DR PROSITEs; PS00135; TRYPSIN_SER_1.
SEQUENCE 1457 AA; 159965 MW; B579F52A5CDD768F CRC64;

Query Match 56.2%; Score 4084; Da 2; Length 1457;
Best Local Similarity 55.2%; Pred. No. 4.8e-168;
Matches 805; Conservative 213; Mismatches 340; Indels 100; Gaps 16;

```

QY 10 FLNCLSLGVSQAAAGHTFGIDYQYTPYKNGKFTGAQNTLYTNKQQLVGSMT 69
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 27 YLAICISFGLIPQAAAGHTFGINQYTRDFANKGKFAAGADIVYNNKKEBLVGSMT 86
QY 70 KAPMIDSVSSNGVAALVENQYIVSAHVAVGYTDVDFGAEKNPQDHFYTKLYKNY 129
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 87 KAPMIDSVSSNGVAALVGDQYIVSAHVAVGYTDVDFGAEKNPQDHFYTKLYKNY 146
QY 130 KCDML-HPEYDDYHNPRLHKFVTEAAPIDMTSNMGSTYSDRKYERVRIGSGRPWRN 188
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 147 KAGTKHRYGQDTHMPRLHKFVTEAEPVMTSYMGKRYTIDQNNYEDRVIRIGAGQYWRN 206
QY 189 DQDQD-----QVAGAHYITAGNTHNQRGANGCYSLGQDVRKAGEVPLPLAGSKD 242
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 207 DEDPPNRBSYHILASYSMLVGNTFAQMGSGGQYVNLASSEKIKSPYGFPLFGSGPFD 266
QY 243 GSGMFTYDAEKOKMLINGILRBGNPESKENGAPOLYKSYF-DEIFERDLTSLYTRAG 301
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 267 GSGMFTYDAEKOKMLINGILRBGNPESKENGAPOLYKSYF-DEIFERDLTSLYTRAG 326
QY 302 NGVYTSISGNDGSGSTYK--SGIPSEIK--ITLANSILPKKOKYHNRYDGNIT 355
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 327 NGKYSFYDDNNGTGKLNKHEHNSLPRKLTQVQLFVNSLSEFAEPYHA--GVVNSY 385
QY 356 SPRLNGETLYFYDQKQSLIFASDINQAGGLYEGNPTWSPNSQNTQAGLHVSNS 415
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 386 RPRLNGENISFIDEKELLITNSINQAGGLYFQSDPTVSPENNETWQGVNISDS 445
QY 416 TVTKVNGVGHDLKSLIGKGLIHQAQKENGKSSISVSGKVIILEQOADQKQAFESIG 475
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 446 TVTKVNGVANDLSLIGKGLIHQAQKENGKSSISVSGVYIILQOADQKQAFESIG 505
QY 476 IVSGRGTVOLNDKQDFTDKFYFGFRGRDLNGLSLFKRIQNTDEGAMIVANTQA 535
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 506 IVSGRGTVOLNANQENPDKLYFGFRGRDLNGLSLFKRIQNTDEGAMIVANTQA 565
QY 536 NWTITGNESTIVLPMGNINIKLDYKKEIAVNGRGETDKNGKNGIMLYKPTERTILL 595
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 566 TVTITGKQDLA-TTGN--NSLDSKKEITANGVFGEDTKTNGKLMVYQPAEDRTLL 623
QY 596 SGGTNLKGDIITQKGLFFSGRPYHAYNHLNKRWSNEMEGIPQETIVMDHMINETFAE 655
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 624 SGGTNLKNMILTOYNGKGLFFSGRPYHAYNHLNKRWSNEMEGIPQETIVMDHMINETFAE 683
QY 656 NFQIKGSAAVSNVSSISBGWTVSNANADTPGVVPOQNTICTRSMDGLTTCOKDILT 715
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 684 NFQIKGSAAVSNVSNVSKWGMHSHNAQVFEVAPHOSHICTRSMDGLTTCOKDILT 743
QY 716 DPKVNSIPKTOIKSINLTNATVANKGLAKNGVTLTHSOFILSNNAIOGN----- 77-
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 744 DPKVNSIPKTOIKSINLTNATVANKGLAKNGVTLTHSOFILSNNAIOGN----- 803
QY 772 -----RLSDNSTATVNNANLNGVATLDS 796
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 804 GNAQATPQATLNGNTSASGNASFWLSDEAVNGSLISGAKNAVSHSLANGVSLADK 863
QY 797 AQSLSKSHFSHOIQDQKFTVYTLNATVYEDDTLQNLTLNNTSTLTSAY----- 849
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 864 AVFHESSRFTQGISGKNTALHLKDSBMTLPSTGLNGLNMDATLTLSAARDAAGA 923
QY 850 -SASGNTPR-----RSLETETPTSAEHRNTLVNGLSQGTFQFTSLFGYSKD 902
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 924 QTSKATDAPRRKRBRKRSLSVTPFTSVSRNTLVNGLNQGFPRMSELFGYSKD 983

```

```

QY 903 KKLNSDAESGYLLSVRNTKEPELLOGLTVESKQNPQSLDKLFTLENDHYCAALRY 962
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 984 KKLNSDAESGYLLSVRNTKEPELLOGLTVESKQNPQSLDKLFTLENDHYCAALRY 1043
QY 963 KLVNDGEFFRLNPIYKQELNDVRAEQARLLEKQVEPFIATQYGEKYSRRARA 1022
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1044 QLRDGEFFRLNPIYKQELNDVRAEQARLLEKQVEPFIATQYGEKYSRRARA 1089
QY 1023 APEPDLTPQOSLNLNLEKQAE-----LTAEQSKAKTKKVSRAVFSDDLDSIFA 1376
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1090 AGRDVA-EKISVAEPHQAQGENVGYQAEBEKRVQDK-----DTALAKREAE 1140
QY 1077 LBALEVIDAPQOESKDRLOEBAKQKQKQISYNSALSSELATVNSMLSVDELD 1136
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1141 TRPATTAFFRARRARBDLPQLOPQOPQOPDILSHYANSGLSEPSATLNSVAVODELD 1203
QY 1137 RLFPDQASAVNTYNAQDRAVDSPAEPAYQOQKTLRQYQALNGLTAVFHSNS 1196
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1201 RYFAEDRRNAVMTSGIRDTYHRSQDFRANYQO-TLRLQIGQKRLSSGKGLFESHNT 1259
QY 1197 DNTFEOYKNAKTLTMMSGPAQYQWGLQGVNVGTGISASQAEQSRKIRKKAINGV 1256
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1260 ENTFDQIGNSARLAKAVFQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQYQY 1319
QY 1257 NASYOFRLOGLQIOPFYGVNRYFIERENYQSEBVRVKTSLAENKYNAGSRVDTTPTD 1316
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1320 QARYRAGSGGIEPHIGATRYFQKDYRENNVNIATPGLAFNRYRAGKADYSKFPQ 1379
QY 1317 NISVAFYFVNVYDVSAVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 1376
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1380 HISITPYLSLYTAAASGKRRVNTAVLADFGKTRSAWENAIKFTLSLHAAAK 1439
QY 1377 GSQLQKQNVGKLGYYW 1394
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1440 GQPLEKQHSAGIKCYRW 1457

```

RESULT 8
Q9X7H1 PRELIMINARY; PRT: 1457 AA.

AC Q9X7H1; 01-NOV-1999 (T-EMBLrel. 12, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE App protein.
GN App.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-15:P1.16;
RA Abdel-Hadi H., Woolridge K.G., Ala Aldeen D.A.;
RT "Identification and characterization of neisserial App."
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ242535; CAB43832.2; -.
DR MEROPS; S06.006; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR005546; Autotransporter.
DR InterPro; IPR001254; Peptidase_91.
DR InterPro; IPR000710; Peptidase_56.
DR InterPro; IPR004899; Peractin.
DR Pfam; PF0197; Autotransporter; 1.
DR Pfam; PF02395; IGAL; 1.
DR Pfam; PF03212; Peractin; 1.
DR PRINTS; PR00921; IGASERPTASE.
DR SMART; SM00020; TYP_SPC; 1.
DR TIGREPMs; TIGR01414; autotrans bar1; 1.
DR PROSITE; PS00135; TRYPSIN_SER_1.

KW Hydrolyase; Protease; Serine protease.
SQ SEQUENCE 1457 AA; 159957 MW; CE62390B6C04B781 CRC64;

Query Match 56.0%; Score 4072; DB 2; Length 1457;
Best Local Similarity 55.0%; Pred. No. 1.6e-167;
Matches 804; Conservative 219; Mismatches 330; Indels 110; Gaps 19;

```

10 PLTACISLGIIVQAMNGHTYFGIDVQYRDFANRKGKFPVQANIKYVNGQOLVGTMT 69
27 YIATLSFGILPQAMNGHTYFGINQYRDFANRKGKFPVQANIKYVNGQOLVGTMT 86
70 KAPMIDFSVSRNGVAALVENDQYIVSAHNGVTTDVDFGABGNPDQHTTYKYKNNY 129
87 KAPMIDFSVSRNGVAALVENDQYIVSAHNGVTTDVDFGABGNPDQHTTYKYKNNY 146
130 KQDNL-HPYEDDYNRLKFTYLAAPIDMTSMNGSTYSRKYPERVAVISGGQOFNRN 188
147 KAGYGHYGGDYHMPRLKFTYLAAPIDMTSMNGSTYSRKYPERVAVISGGQOFNRN 206
189 DDDKGD-----QVAGAYHYLTAGNTHQKQAGNGSYLGADVERKAGEYGPLTLAGSKD 242
207 DEDERNRBSSTYHLSYSLMGNTFPAQNSGGGTVMGSEKIKHPYGFPLTGGSPGD 266
243 SSGPMFYDAKQKMLINGILRBNPFEKENGQOLVKRSTF-DEIFRDLATSLYTRAG 301
267 SSGPMFYDAKQKMLINGILRBNPFEKENGQOLVKRSTF-DEIFRDLATSLYTRAG 326
302 NGYITISGNDGQSSITQK---SGIPSEIK---ITLANNSLDKKQKXNRPYDQPNLY 355
327 NGKYSFNDNNKSTKINAKENHNSLPNLRKRTYQVLPNVSLETAKEVYTHAA-GEVNSY 385
356 SPLANNGETLYFMDQXGSLIPASDINGAGLTFYBGNFTVPSNNSQWQAGIHVENS 415
386 RPLANKGENISFTIDEGGELLISNINOGAGGLYFGQDPTVSPENNERWQAGVHSEDS 445
416 TYTWKNGVEHDLRSKIGKGLTHVQAKSENGKSTISVQDKVTLLEQADDDQAKFSEIG 475
446 TYTWKNGVANDRLSKIGKGLTHVQAKSENGKSTISVQDKVTLLEQADDDQAKFSEIG 505
476 LVSGRGTQVLDNDQKQPTDKEFYRGRGRDLNNGHSTTPRRIQNDDEAMVYNNHTQAA 535
506 LVSGRGTQVLDNDQKQPTDKEFYRGRGRDLNNGHSTTPRRIQNDDEAMVYNNHTQAA 565
536 NVYITGENESIVLPNGNNINFLDYRKEIAVNGFSETDKNKNGRNLILYKPTTEETLL 595
566 TVYITGNKDLA-TTGN-NSLDSKEELAYNGFEBQOTTKINGRLNLVYQAAEDDTLL 623
596 SSGTNLKGDIQYTKGLFSGRPTPHAVNHLNKRWSMEGIPQSEITWDHDMNIRPKAE 655
624 SSGTNLNGNITQYTKGLFSGRPTPHAVNHLNKRWSMEGIPQSEITWDHDMNIRPKAE 683
656 NFOIKGSAVYSRVSSIEGNWTVSNANATFCYVYQOQNTICTSDWYGLTQCKYDLT 715
684 NFOIKGSAVYSRVSSIEGNWTVSNANATFCYVYQOQNTICTSDWYGLTQCKYDLT 743
716 DTKVINSIKPTQINSINLTDMATAVYKGL----- 745
744 DDKVINSIKPTQINSINLTDMATAVYKGL----- 783
746 -----AKINGNVTLNHSQFTLSNNATQIGNIRLSNGSTAYTDNANLNQVHLLTS 796
804 GNAQATFNQATLNGNTSASGNASFNLSHDAVNGSLTSLGNAVANSALNANVSLADK 863
797 AQSIAKSHFHSIQCKGKTYLLEAATWMSQTLTONTLANSSTITLNSAY----- 849
864 AVFHESSSRFTQISGGKDTALHLKQSEMTLPSGELGNANLNATITLNSAVERHAKA 923
850 -SASNNTPR-----RSLETEFTTSAEHRFNTLVNKLSSGQFTQTSLSFGYKD 902
924 QTSATDAPARRRRSRSSRLSVTPPTVSSEFNTLVNKLSSGQFTQTSLSFGYKD 983
903 KTKLSNDAGDYILSVRNTGSEPTLEQTLTVESNDQPSLCKKFTLENDHNDAGLRY 962

```

```

DB 964 KTLAASSGCTYTAAVNTNGNPASILEQITVYBGKDKKPLISENLNPTQNEHVDAKMY 1043
QY KLVANDGEFFRLNPIKQELNDLVRAE---QABR-----TLBAQVEPTAQTQ-GE 1011
DB 1044 QLRNDGEFFRLNPIKQELNDLVRAE---QABR-----TLBAQVEPTAQTQ-GE 1103
QY 1012 PKVSRRAAPAAFPDTLPDQSLNLEAKQELTAEQSKATKPKSKAVSDPDL 1071
DB 1104 P-----ARQAGEHV---GIMQAEEEK--RYQADKDTALAKQREAEPTPTAFA 1149
QY 1072 QSLFLEALAEVIDAPQSEKDLQAEAKQKQKOLISRYNSALSELATVNSMLSV 1131
DB 1150 -----BARARADLPQIQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 1195
QY 1132 QDELDELFDQASAVMTNIAODKRRYDSDAFRAYOQKTNLQISVQPLANRIGAVF 1191
DB 1196 QDELDELFDQASAVMTNIAODKRRYDSDAFRAYOQKTNLQISVQPLANRIGAVF 1254
QY 1192 SHSRSDNTFDSQVCKHATLTMGSGFAQYQMGDLQFQVNVGTGISAKMAEBSQKLRKA 1251
DB 1255 SHNRTEFTPDGIGNSARLHAGVFCQYIGIRFYGISAGAGFSQSLSDGIGKIRRV 1314
QY 1252 INGVNASVQFRLQGLQIPYFQVNGYFTEERNYQSEVRYKTPS-APRRYAGIRVDY 1311
DB 1315 LHYGIQKTRAGCGGGLBPGIGATRYFQKDYKRYEYVNTATPGLAEKRYAGIKADYS 1374
QY 1312 FTPTDNTSVYEFYFNVYDVSNANVOTVNLTVLQPPGRYQKQVGLKABTLHFQISAF 1371
DB 1375 FKPAQHSITPPLSITDASGKVRVNTVYLAQDEGKTRSEMGVNAEIKGFTLSLH 1434
QY 1372 ISKSQSQQLKQOQNGVVLGYRM 1394
DB 1435 AAAGKQPLRQHSAGIKLGYRM 1457

```

RESULT 9
Q9F324 PRELIMINARY; PRT; 1454 AA.
AC Q9F324;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, last annotation update)
DE Adhesin penetration protein.
GN App.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=24181;
RA Ala-Aldeen D.A.;
RL Submitted (0CT-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=24181;
RA Abdel Hadi H., Woodridge K.G., Robinson K., Ala Aldeen D.A.;
RT "Identification and Characterization of App: an Immunogenic
Autotransporter Protein of Neisseria meningitidis";
RL Submitted (0CT-2000) to the EMBL/GenBank/DBJ databases.
EMBL: A0296277; CAC14671.1; --
DR MEROPS; S06.006; --
DR GO; GO:008233; F:peptidase activity; IEA.
DR GO; GO:004295; F:trypsin activity; IEA.
DR GO; GO:006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR00546; Autotransporter.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR000710; Peptidase_S6.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02395; IGA1; 1.
DR Pfam; PF03212; Pertactin; 1.

SQ SEQUENCE 507 AA; 56170 MW; 665E365FEFE273F1 CRC64;
 Query Match 34.7%; Score 2523; DB 2; Length 507;
 Best Local Similarity 92.2%; Pred. No. 2.8e-101;
 Matches 471; Conservative 18; Mismatches 18; Indels 4; Gaps 1;

QY 70 KAPMIDFSVSRNGVAAALVENQYIVSAHNVGYDVDFABGANNPDRHFFTYKIVKRNYY 129
 DB 1 KAPMIDFSVSRNGVAAALVENQYIVSAHNVGYDVDFABGANNPDRHFFTYKIVKRNYY 60
 QY 130 KCONLHPYEDDYHNPRLHKEVTEAPIDMTSMNNGSTYSDETKYPERVIRIGSGQFWEND 189
 DB 61 KCONLHPYEDDYHNPRLHKEVTEAPIDMTSMNNGSTYSDETKYPERVIRIGSGQFWEND 120
 QY 190 QDKGQVAGAHYHILTAANTHNOGAGNGSYLGGDVRAKAGYGPPLIAGSGDSSGSPMT 249
 DB 121 QDKGQVAGAHYHILTAANTHNOGAGNGSYLGGDVRAKAGYGPPLIAGSGDSSGSPMT 180
 QY 250 YDAERKQWMLINGILREGNPEKENGFOIVKRSYFDEIFERDLHTSLYTRAGNGVYITISG 309
 DB 181 YDAERKQWMLINGILREGNPEKENGFOIVKRSYFDEIFERDLHTSLYTRAGNGVYITISG 240
 QY 310 NDNGGSIITQSGIPESEIKITLANNSLPLKEXKVHNPRIYDGNITSPRLNNGETLYTMD 369
 DB 241 NDNGGSIITQSGIPESEIKITLANNSLPLKEXKVHNPRIYDGNITSPRLNNGETLYTMD 300
 QY 370 QKQGSILFASDINGAGGLYFEGNFTVSPNSNQTWOGAGIHVSNSYTWKVGVEHDL 429
 DB 301 QKQGSILFASDINGAGGLYFEGNFTVSPNSNQTWOGAGIHVSNSYTWKVGVEHDL 360
 QY 430 SKIGKGLHVOAKGKNGSISVGDGKYLBOQADDOGNKQAFSEIGLVSGRGVOLNDK 489
 DB 361 SKIGKGLHVOAKGKNGSISVGDGKYLBOQADDOGNKQAFSEIGLVSGRGVOLNDK 420
 QY 490 QPDTDXEYFGRGRLDNGSLTFKRIQNTDEGAMIVNHTTQAAVNTTIGNESTVLPN 549
 DB 421 QPDTDXEYFGRGRLDNGSLTFKRIQNTDEGAMIVNHTTQAAVNTTIGNESTVLPN 476
 QY 550 GNNINKLDYRKELIANNVNGFETDKNGKRL 580
 DB 477 NDILKQLTNRKDIAFNMGFETDKNGKRL 507

RESULT 11
 Q8KRG5 PRELIMINARY; PRT; 505 AA.
 ID Q8KRG5
 AC Q8KRG5;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hap (Fragment)
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HK274, and HK75.
 RA Kilian M., Poulsen K., Lomholt H.;
 RT "Evolution of the paralogous hap and iga genes in Haemophilus
 RT influenzae: evidence for a conserved hap pseudogene associated with
 RT microcolony formation in the recently diverged Haemophilus aegyptius
 RT and H. influenzae aegyptius";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF517141; AAM74927.1;
 DR EMBL; AF517142; AAM74928.1;
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR000710; Peptidase_S6.
 DR Pfam; PF02395; IGA1; 1.
 DR PRINTS; PR00921; IGASERPTASE.

DR PROSITE; PS00135; TRYPsin_SER; 1.
 FT NON TER 1
 FT NON TER 505
 SQ SEQUENCE 505 AA; 55622 MW; ABAC61F4CD54494 CRC64;
 Query Match 29.3%; Score 2128; DB 2; Length 505;
 Best Local Similarity 77.5%; Pred. No. 3e-84;
 Matches 396; Conservative 49; Mismatches 60; Indels 6; Gaps 2;

QY 70 KAPMIDFSVSRNGVAAALVENQYIVSAHNVGYDVDFABGANNPDRHFFTYKIVKRNYY 129
 DB 1 KAPMIDFSVSRNGVAAALVENQYIVSAHNVGYDVDFABGANNPDRHFFTYKIVKRNYY 60
 QY 130 KCONLHPYEDDYHNPRLHKEVTEAPIDMTSMNNGSTYSDETKYPERVIRIGSGQFWEND 189
 DB 61 KCONLHPYEDDYHNPRLHKEVTEAPIDMTSMNNGSTYSDETKYPERVIRIGSGQFWEND 120
 QY 190 QDKGQVAGAHYHILTAANTHNOGAGNGSYLGGDVRAKAGYGPPLIAGSGDSSGSPMT 249
 DB 121 QDKGQVAGAHYHILTAANTHNOGAGNGSYLGGDVRAKAGYGPPLIAGSGDSSGSPMT 180
 QY 250 YDAERKQWMLINGILREGNPEKENGFOIVKRSYFDEIFERDLHTSLYTRAGNGVYITISG 309
 DB 181 YDAERKQWMLINGILREGNPEKENGFOIVKRSYFDEIFERDLHTSLYTRAGNGVYITISG 240
 QY 310 NDNGGSIITQSGIPESEIKITLANNSLPLKEXKVHNPRIYDGNITSPRLNNGETLYTMD 369
 DB 241 NDNGGSIITQSGIPESEIKITLANNSLPLKEXKVHNPRIYDGNITSPRLNNGETLYTMD 300
 QY 370 QKQGSILFASDINGAGGLYFEGNFTVSPNSNQTWOGAGIHVSNSYTWKVGVEHDL 429
 DB 301 QKQGSILFASDINGAGGLYFEGNFTVSPNSNQTWOGAGIHVSNSYTWKVGVEHDL 360
 QY 430 SKIGKGLHVOAKGKNGSISVGDGKYLBOQADDOGNKQAFSEIGLVSGRGVOLNDK 489
 DB 361 SKIGKGLHVOAKGKNGSISVGDGKYLBOQADDOGNKQAFSEIGLVSGRGVOLNDK 420
 QY 490 QPDTDXEYFGRGRLDNGSLTFKRIQNTDEGAMIVNHTTQAAVNTTIGNESTVLPN 549
 DB 421 QPDTDXEYFGRGRLDNGSLTFKRIQNTDEGAMIVNHTTQAAVNTTIGNESTVLPN 476
 QY 550 GNNINKLDYRKELIANNVNGFETDKNGKRL 580
 DB 475 NDILKQLTNRKDIAFNMGFETDKNGKRL 505

RESULT 12
 Q8KSC0 PRELIMINARY; PRT; 510 AA.
 ID Q8KSC0
 AC Q8KSC0;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hap (Fragment)
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HK61;
 RA Kilian M., Poulsen K., Lomholt H.;
 RT "Evolution of the paralogous hap and iga genes in Haemophilus
 RT influenzae: evidence for a conserved hap pseudogene associated with
 RT microcolony formation in the recently diverged Haemophilus aegyptius
 RT and H. influenzae aegyptius";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF517151; AAM74935.1;
 DR EMBL; AF517152; AAM74936.1;
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR000710; Peptidase_S6.

Pfam: PF02395; IGAL; 1.
 DR PRINTS; PR000921; IGASERPHASE.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 FT MON TER 1 1
 FT NON TER 510 510
 SQ SEQUENCE 510 AA; 56795 MW; C5746C0C18949E4D CRC64;

Query Match 28.7%; Score 2087.5; DB 2; Length 510;
 Best Local Similarity 76.2%; Pred. No. 1,76-82;
 Matches 390; Conservative 51; Mismatches 68; Indels 3; Gaps 2;

QY 70 KAPMIDFSVSRNGVAALVENQYIVSAHNGVTDVDFGAAGNPDQHFYKIVKNNY 129
 DB 1 KAPMIDFSVSRNGVAALVENQYIVSAHNGVTDVDFGAAGNPDQHFYKIVKNNY 60
 QY 130 KKNLHPEYEDYHNPRLHKEVTEAPIDMTSNMGSTYSDRKTYPERVRIOSGRCPKRD 189
 DB 61 NHDKAKHYLDYHNPRLHKEVTEAPIDMTSHDCKYANKKPKYPERVVGSGDYMDDD 120
 QY 129 QDKGQYAAQVHYLTAGTNHNGAGNGSYLGQVYRKAGYGPPLIAGSKDGSMPFI 249
 DB 121 QNNRTYLSGGNTYLTGGNTYNOGSGDGSISYVAGDIRKVDYGPPLIAGSKDGSMPFI 180
 QY 250 YDAEKOKMLINGILREGNPFEGKNGFQVLRKSYFDEIFERDLHTSLYTRAGNGVYITSG 309
 DB 181 YDAETQKMLINGILREGNPFEGKNGFQVLRKSYFDEIFERDLHTSLYTRAGNGVYITSG 240
 QY 310 NDNGGGSITQKSGITSEIKITLAMSILPKKDKQTHNRYNRYGPIYTSPLNNGELTLYMD 369
 DB 241 SDGGIGVVTYSKIKRPFEE--IPANNKLIKEDKDTYNNRYNGPIYTSPLNNGKXLYFGD 238
 QY 370 QKQGSILFASDINGQAGLYEFGNFTVSPNSNOTQAGHVSNSITVMKYNVGEHDL 429
 DB 299 EELGSITLITLDGQAGLYEFGNFTVSPNSNOTQAGHVSNSITVMKYNVGEHDL 358
 QY 430 SKIKGTLHVAQKNGKSGISVGDGKYLLEQADQKQKQAFSEIGLVSGKGTQVLDNDK 489
 DB 359 SKIKGTLHVAQKNGKSGISVGDGKYLLEQADQKQKQAFSEIGLVSGKGTQVLDNDK 418
 QY 490 QPDDTKYFFGFRGRLDLNGSLTPKRIQNTDEGAMVYNNHTQQAANVTITNESIVLP 548
 DB 419 QPDDTKYFFGFRGRLDLNGSLTPKRIQNTDEGAMVYNNHTQQAANVTITNESIVLP 478
 QY 549 NGNNINKLDYRKEIAYNGFGETDKNGKNGRL 580
 DB 479 NGNNINKLDYRKEIAYNGFGETDKNGKNGRL 510

RESULT 13

QKRSB9 PRELIMINARY; PRT; 526 AA.
 ID QKRSB9
 AC QKRSB9;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hap (Fragment).
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=HK715;
 RA "Kilian M., Poulsen K., Lomholt H.;
 RT "Evolution of the paralogous hap and iga genes in Haemophilus
 RT influenzae: evidence for a conserved hap pseudogene associated with
 RT microcolony formation in the recently diverged Haemophilus aegyptius
 RT and H. influenzae aegyptius.";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF517152; AAM74936.1; -
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0005508; F:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR009003; Cys_Set_trypsin.

DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR000710; Peptidase_S6.
 DR Pfam; PF02395; IGAL; 1.
 DR PRINTS; PR000921; IGASERPHASE.
 DR PROSITE; PS00135; TRYPsin_SER; 1.
 FT MON TER 1 1
 FT NON TER 526 526
 SQ SEQUENCE 526 AA; 58159 MW; 4B77316D9E1B4F74 CRC64;

Query Match 25.3%; Score 1839.5; DB 2; Length 526;
 Best Local Similarity 66.7%; Pred. No. 96-72;
 Matches 355; Conservative 52; Mismatches 98; Indels 27; Gaps 6;

QY 70 KAPMIDFSVSRNGVAALVENQYIVSAHNGVTDVDFGAAGNPDQHFYKIVKNNY 129
 DB 1 KAPMIDFSVSRNGVAALVENQYIVSAHNGVTDVDFGAAGNPDQHFYKIVKNNY 60
 QY 130 KKNLHPEYEDYHNPRLHKEVTEAPIDMTSNMGSTYSDRKTYPERVRIOSGRCPKRD 189
 DB 61 KPGKPNFYHGHYHNPRLHKEVTEAPIDMTSNMGSTYSDRKTYPERVRIOSGRCPKRD 120
 QY 190 QDK-----GDQVAGVHYLTAGTNHNGAGNGSYLGQVYRKAGYGPPLIAGSKDGS 243
 DB 121 KQSEYNAVSSTDISGATNYLMAENTHTQSGDNGTYHFGSNVIRPNRYGPIIAGAGDS 180
 QY 244 GSPMFIYDAERQKMLINGILREGNPFEGKNGFQVLRKSYF-----DEIFERDLHT 294
 DB 181 GSPMFIYDAERQKMLINGILREGNPFEGKNGFQVLRKSYF-----DEIFERDLHT 236
 QY 295 SLYTRAGNGVYITSGNDNGGSLT-----QKSGITSEIKITLAMSILPKKDKQ--KYNNRY 349
 DB 237 --YIPISNHSYFSSNNDGKYLITRPSKQSKASQVGYKLNPSELETKERAKAA 294
 QY 350 DGPNTSPRLNNGETLYFDQKQGSILFASDINGQAGLYEFGNFTVSPNSNOTQAGI 409
 DB 295 PGNITQPMRHSKNTITFGDRGTGLTENTINGQAGLYEFGNFTVSPNSNOTQAGI 354
 QY 410 HVSNSITVMKYNVGEHDLRLSKIKGTLHVAQKNGKSGISVGDGKYLLEQADQKQ 469
 DB 355 HVSNSITVMKYNVGEHDLRLSKIKGTLHVAQKNGKSGISVGDGKYLLEQADQKQ 414
 QY 470 ABEFISLVSGSGTQVLDNDKQPTDKYFFGFRGRLDLNGSLTPKRIQNTDEGAMVY 529
 DB 415 ABEFISLVSGSGTQVLDNDKQPTDKYFFGFRGRLDLNGSLTPKRIQNTDEGAMVY 474
 QY 530 NTQQAANVTITNESIVLP--NGNNINKLDYRKEIAYNGFGETDKNGKNGRL 580
 DB 475 NTQQAANVTITNESIVLP--NGNNINKLDYRKEIAYNGFGETDKNGKNGRL 526

RESULT 14

QKRN0 PRELIMINARY; PRT; 514 AA.
 ID QKRN0
 AC QKRN0;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hap (Fragment).
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=HK284; and HK292;
 RA "Kilian M., Poulsen K., Lomholt H.;
 RT "Evolution of the paralogous hap and iga genes in Haemophilus
 RT influenzae: evidence for a conserved hap pseudogene associated with
 RT microcolony formation in the recently diverged Haemophilus aegyptius
 RT and H. influenzae aegyptius.";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF517143; AAM74929.1; -
 DR EMBL; AF517145; AAM74931.1; -

DR GO: 0004295; F: trypsin activity; IEA.
 DR GO: 0006508; P: proteolysis and peptidolysis; IEA.
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR00710; Peptidase_S6.
 DR Pfam: PF02395; IGAL; 1.
 DR PRINTS: PRO0921; IGASERPTASE.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 FT NON_TER 1
 FT NON_TER 514
 SQ SEQUENCE 514 AA; 56562 MW; 8848BDDPFOA313 CRC64;
 Query Match 23.7%; Score 1724.5; DB 2; Length 514;
 Best Local Similarity 65.8%; Pred. No. 8e-67;
 Matches 342; Conservative 53; Mismatches 110; Indels 15; Gaps 8;

QY 70 KAPMIDFSVSNAGVALVENQYIVSAHNVGYTVDVFGAEGNNDQHRFTYKIVERNY 129
 DB 1 KAPMIDFSVSNAGVALVENQYIVSAHNVGYTVDVFGAEGNNDQHRFTYKIVERNY 60
 QY 130 KQDNHLYEDDTHPRHKKVTEAPIMTNNMGSTYSPTKTYSERVAGSGRPFWRMD 189
 DB 61 KPGQ--XDGTHPRHKKVTEAPIMTNNMGSTYSPTKTYSERVAGSGRPFWRMD 117
 QY 190 --ODKGDQ---VAGAYHYLLTAGTNHORGANGSYLGGDYVXKAGEYPLTAGSKGDS 243
 DB 118 EQAKADAVSSWLAGAIVNKLKAGNTHVOTGNGTVDLSGNLTKXNHGFLPIAGSGDS 177
 QY 244 GSDMFTYDAEKQKMLINGILREGNPEKENGFPOLYKSYF-DEIFERDLHTSLYTRASN 302
 DB 178 GSDMFTYDAEKQKMLINGILREGNPEKENGFPOLYKSYF-DEIFERDLHTSLYTRASN 237
 QY 303 GYTTISGNDNGCSITQKSGIPKILITAMSLP-EKMDKYNHRYDGPVTSPLNNG 362
 DB 238 KHYLFYANDNGIGTVL-KTEDSTSTTVKLEPITLSEGVKRY--ARGANNFTKPKLDNG 294
 QY 363 ELYEYNDKQKSLIFASDINQAGALYFEKNFTVS--PNSNOTWQAGIHVSENSTYTWY 421
 DB 295 ELSIFIDQKGEILFTNSVWQAGALYFEKGFVDSTANPRMDIWQAGSISTEDTYYWY 354
 QY 422 NGVEHDLKIKIGKGLTHVQAKGKNGKSSIVQDGYVILEQADDOQKQASEICLYSGG 481
 DB 355 KNEPGRILSKIGLTLVNGTGKLGNTISVGVGVILLDQKADIDGKQAFKXVGVISGR 414
 QY 482 TYQANDKQFPTDKYFGRGRLDNGSLTFKRIQNTQBGAMVNNHTQAAVNTTG 541
 DB 415 TYQINSADQVDNNIYFGRGRLDNGSLTFKRIQNTQBGAMVNNHTQAAVNTTG 474
 QY 542 NESIVLP-NGANNINLDYRKEIAYNGMFGEDTKXKNGRL 580
 DB 475 NESIVLP-NGANNINLDYRKEIAYNGMFGEDTKXKNGRL 514

RESULT 15
 Q9JXK3 PRELIMINARY; PRT; 1431 AA.
 AC Q9JXK3:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Serine-type peptidase.
 GN NMB1998.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,

RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
 RA Colton M.D., Uitterlinden T.R., Khouri H., Qin H., Yamashiro Y.,
 RA Gill J., Scarlato V., Maignani V., Pizzo M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.W., Moxon E.R., Rappelli R., Venter C.C.,
 RA "Complete genome sequence of Neisseria meningitidis serogroup B strain
 RT MC58.";
 RL Science 247:1809-1815 (2000).
 DR EMBL: AF002549; AAF42325.1; -.
 DR PIR: A81018; A81018.
 DR TIGR: AMB1998; -.
 DR GO: 0004295; F: trypsin activity; IEA.
 DR GO: 0006508; P: proteolysis and peptidolysis; IEA.
 DR InterPro: IPR006315; Autotransporter.
 DR InterPro: IPR005546; Autotransporter.
 DR InterPro: IPR009003; Cys_Ser_trypsin.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR00710; Peptidase_S6.
 DR InterPro: IPR04899; Peptidase_S6.
 DR Pfam: PF03197; Autotransporter; 1.
 DR Pfam: PF02395; IGAL; 1.
 DR PRINTS: PRO0921; IGASERPTASE.
 DR SMART: SM00020; ITPY-SPEC; 1.
 DR TIGRPFAM: TIGR01414; autotrans_bar1; 1.
 DR Complete proteome.
 SQ SEQUENCE 1431 AA; 157645 MW; 6801F0E209C24EFC CRC64;

Query Match 23.0%; Score 1676; DB 16; Length 1431;
 Best Local Similarity 29.7%; Pred. No. 3.8e-64;
 Matches 470; Conservative 229; Mismatches 524; Indels 362; Gaps 50;

QY 10 FTACISISGI--VSQAMAGHTYFGIDQYRDEAKNKFTVGAQNTKYNKQGLVGS 67
 DB 9 FCSVSTLGLFANSPAYSSIVRNDVDQYFRDEAKNKGFVGAQNSISIDKQGLKGV 68
 QY 68 MKKAPMIDFSVSR-NGVALVENQYIVSAHNVGYTVDVFGAEGNNDQHRFTYKIV 126
 DB 69 LINGIMPEDFVSRNROFALITLVHPQVNSVKNVGVSGSIOGNTQMPBEQOAYELVSR 128
 QY 127 NNYKKNNLHYPED-DYHNPRLKHTVTEAPIDMTS-----NMNGSTYCDRTKYERVR 178
 DB 129 N-----PDPYDHLPRKLVTEISPTALSSVPLKNGQKRAKAYLDTDRPFYPR 180
 QY 179 ISSGRQFWRMDQKQDVAGAYHYLLTAGTNHORGANGSYLGGDYVXKAGEYPLTAG 238
 DB 181 LSGTQYVRKADGTRRTAPAYQVLTGTPPLKVLGQVHGLVGGSLTQ----PLNTYA 236
 QY 239 SKGDSGSMFTYDAEKQKMLINGILREGNPEKENGFPOLYKSYFDEIFERDLHTSL 296
 DB 237 IAGDGSPLFAFDHNRNFWLAGVL--STYAGFDFPKYIVTPPEFIRSTRQYETRL 293
 QY 297 YTRAGNGVYTESGNDNGQGSITQKSG-PSEIKITLNMSLP-EKMDKYNHRYDGPVTS 356
 DB 294 --DVELITNELIKRDNGNGLTQ-GLAKERTLPIANPSL-APQNDRAHNPSSDA----- 344
 QY 357 PLNNGETLYE--MDQKQSLIFASDINQAGALYFEKNFTVSNSQVWQAGIHVSE 413
 DB 345 -----GKTLISSRPNK--TLMLADNINQAGALQFDSNFTV--VGKHTWQAGAVTAD 396
 QY 414 NSTVTWKVGVHEDRLSKIGKGLTHVQAKGKNGKSSIVQDGYVILEQADDOQKQASE 473
 DB 397 GKRVFQVNSKRGRLSKIGAGTLIANQGINQSDISIEGTGVLLQKQASDSCQKAFNQ 456
 QY 474 IGVSARGVQVNDKQFPTDKYFGRGRLDNGSLTFKRIQNTQBGAMVNNHTQAAVNTTG 533
 DB 457 VGTISRGVAVLADSOQIKPEMLYFGRGRLDNGSLTFKRIQNTQBGAMVNNHTQAAVNTTG 516
 QY 534 AANVTITANE-----SIYPNKR----- 551
 DB 517 AATLITLTGNPILSPHEVWQVNGNRPQNAAYEYINPRNRRTDYFLXPQGNPREPP 576
 QY 552 -----NINLDRKELA-----YNGMFGEDTKXKNGRL 571

```

Db 577 LMKONSTWOFIGNRQQAEOVAQENARBDLITPGYIGENRQGTGAAPSYSTNMEA 636
QY 572 -DNXK-----NGRLNLYKPTEDRTLLSGTNLKGDTOTKGLPPSG 616
Db 637 IEKTRHANAVYKAPRYNGALNLYPRKRTDSTLLNGMNLNGEVLIGGMIVSG 696
QY 617 RPTTHANHLNKRNSMEGIPQGEIYVHDHMINRTFKENFOIKGSAAVS-RNVSSIEG 675
Db 697 RPYHAJDHQAAR-----BPVLENEWTDSSFKAAEFTLNARLTAGRNPAHDG 746
QY 676 NMTSSNANATPGVNPNOONTICTRSDWGLTTCOKVDLTUTKYINSIPKQINGSINTL 735
Db 747 DITAYDLSGIDJFTOG-KTPECTRYSHSGSTHCTPNAYLKAENYALPATQVRGDITIN 805
QY 736 DNATANYGLAKINGVTLITNHSQFTLSNNAOIQNIRLSDNSTATVDNANINGVELTD 795
Db 806 DRS-----ELRL-----GKALYG 819
QY 796 SNOJSLNKSHPHQIQGKGTYYLE-NATWTPSDTLQNLTLNNGSTTLNSAYSASN 854
Db 820 S-----IRAGDITAVREADSNWTLISOSHTGALTLIDGAQITLNDPFAANTH 866
QY 855 NTPRRSLETTETPTSABHRFNTLVNGKLSGGTFOF-TSSLFGYKSDKLSNDAGD 913
Db 867 N-----NRFNTLVNGTLDGEGTFRFLTGIVAKQNAAPPLKLGDSRGA 909
QY 914 YILSVRNTGKEPTELEQTLTVESKNOPLSDKLTLENDVADGALRYLVRNDEFR 973
Db 910 FOIHVKTGOEPQTESLAVSL-NEKISHQAFPTLQNGYADIGAFRYILKRNNNGYS 967
QY 974 HNPKEQELNDLVRAE-----QAERLBAKOVEPTAKTGTGEPKVSRAARAAPPTL 1028
Db 968 YNPLEKALQTEATRABHRNOQANQOQTDI-----SRQVQHDSD 1009
QY 1029 POSILNABAKQAEIT--AETOXSNAKTKYRSKRAVFSDDPLD-----QSL-----F 1375
Db 1010 ATRQALQAMQNSQTEIARIDSGVYLSKQKQ-----TDPLTGILTRAQNLCAQGY 1061
QY 1076 ALEAALVETDAPQOS-----EKORLAQBEAKOKOKO----- 1108
Db 1062 SADICQVAKAJTNDLITFETELDTYIERVEMASELDRKQGGDAQAVETARRAYLNA 1121
QY 1109 -----LISRYNSALSELGATVNSMLSVODELDRLEFVDA 1143
Db 1122 IARLSRQJHSIKTGVAGIRMPYLAHLISRSANTAVSEQAAYRTGROQAGRIDRHLDPQ 1181
QY 1144 QSAVWNTLAQDKERYSDAFRAYOQOKTNLRQIGVQKALANG-RIGAVFSHSSDNTPE 1202
Db 1182 QONWLEVTGQDTHSGTHRPY-QQTTNAHIGIQGTIDRLSVGTILTDERTNRPDE 1240
QY 1203 QV-----KNHKLITMNSGPAQYQWMDLOFGVNTGTGISAKAEQSKIHRYKAINGVN 1257
Db 1241 GVSABNRNSGAHL-----FVKGENALFPAADIGYNSRTFTDYDGAAVRRHAMDAGIN 1295
QY 1258 ASYQFLGQGLQIOPYGVNRYFIEEN--YOSEEVKTPSLAFNRYNAGIVDYTFP 1314
Db 1296 TGIKIDTG-INTRPYAGIR--INSNGKRYVUDGALINSRQIQTTMHGIRLDKT-VE 1350
QY 1315 TNSISVKEPFVNVYDV-SNANVQTTVN-LTVLQOPFG--RYWQKEVGLKAEILHFOIS 1369
Db 1351 LQOAKLTPAFSSDYVHPRONSGSALSVNDRILLQAAHGLTHTLIQIDAGYKQ---WNAK 1406
QY 1370 APTSKSGSGLGRKQNVGYKGYRW 1394
Db 1407 LHAAYGDSNTARRKQAGIKIGYNW 1431

```

1	7273	100.0	1394	2	AAR92766	Aar92766	Adnestrion
2	7273	100.0	1394	6	ADA09343	Ada09343	Haemophil
3	7265.5	99.6	1395	6	ADA09348	Ada09348	Haemophil
4	613.5	84.1	1434	6	ADA09350	Ada09350	Haemophil
5	5827.5	80.1	1411	6	ADA09358	Ada09358	Haemophil
6	5807	79.8	1436	6	ADA09354	Ada09354	Haemophil
7	5303.5	72.9	1391	6	ADA09356	Ada09356	Haemophil
8	5297.5	72.8	1391	6	ADA09352	Ada09352	Haemophil
9	4178	57.4	1449	6	ABU37787	Abu37787	Protein e
10	4090.5	56.2	1454	3	AAV56621	Av56621	Neisseria
11	4085	56.2	1457	3	AAH58592	Aah58592	N. mening
12	4085	56.2	1457	6	AAU27556	Aau27556	Neisseria
13	4085	56.2	1457	4	AAU08601	Aau08601	Neisseria
14	4083	56.1	1457	3	AAV56622	Av56622	Neisseria
15	4082	56.1	1457	3	AAH25662	Aah25662	N. mening
16	4076	56.0	1457	2	AAV38823	Av38823	Neisseria
17	4072.5	56.0	1458	6	ABP77279	Abp77279	N. gonorr
18	4072.5	56.0	1458	6	ABU37103	Abu37103	Protein e
19	4059	55.9	1449	2	AAV38824	Av38824	Neisseria
20	4061.5	55.8	1468	2	AAV38825	Av38825	Neisseria
21	4044.5	55.6	1439	4	AAU27557	Aau27557	Neisseria
22	3409.5	46.9	1178	6	AAU08603	Aau08603	Neisseria
23	3354.5	46.1	1136	6	AAU08605	Aau08605	Neisseria
24	2941	40.4	956	6	AAU08602	Aau08602	Neisseria
25	2886	39.7	914	6	AAU08604	Aau08604	Neisseria

26	2334.5	32.1	998	2	AA388822
27	1676	23.0	1431	3	AA755664
28	1673.5	23.0	1541	2	AA072704
29	1673.5	23.0	1541	6	AD009344
30	1662.5	22.9	1545	6	AD009345
31	1647	22.6	1593	2	AA277705
32	1615.5	22.2	1593	6	ABJ37338
33	1609	22.1	1694	6	ABH30453
34	1607.5	22.1	1565	6	ABF78087
35	1605	22.1	1702	6	AD009346
36	1551	21.3	1748	6	AD009347
37	156.5	21.1	1773	6	ABJ37893
38	1144	15.7	501	6	ABU08608
39	1089	15.0	1371	4	ABBS2745
40	1051	14.5	1376	4	ABBS2592
41	675.5	9.3	279	6	ABU08607
42	468.5	6.4	222	6	ABU08606
43	466	6.4	71	1	AAE01315
44	346.5	4.8	73	6	AD009385
45	344.5	4.7	1569	4	AG09842
					Agg98842 E. coli 9

ALIGNMENTS

RESULT 1
AAB02760

ID	AA
AA92768	standard; protein; 1394 AA

AC AAR92768 ;
YY

DT	01-JUN-1996	(first entry)
XY		

DE	Adhesion and penetration protein.
YV	

KW Adhesion and penetration protein; nap gene; signal peptide; protease;
KW helmer domain: pore: outer membrane protein: beta-domain: secretion;

KW diagnostic: immunosay.

XX Haemophilus influenzae -

XX	Yav	Location/Mailifiers
FH		

FT	Peptide	1. .25	/note= "Signal pept
FT			

Protein	26. .959
PT	{note= "Secreted 110-kDa protease fragment"
PT	

FT	Peptide	27.33	/note= "Mature protein N-terminal peptide"
FT			

FT	Domain	241. .248	/note="Putative catalytic domain"
FT			

PT	Active-site	243
PT	/note= "Active site serine residue"	

PT	Region	698	/note= "Conserved cysteine residue"
PT			

FT	Region	/note= "Conserved cysteine residue"
FT		

Protein	pI	Mr	Accession	Database
960.1394	960.1394	960.1394	960.1394	960.1394
45-kDa outer membrane protein	45-kDa outer membrane protein	45-kDa outer membrane protein	45-kDa outer membrane protein	45-kDa outer membrane protein

FT	960	Site for cleavage of secreted proctase from outer
FT		/note=
Cleavage-site		

Et XX

PN MO9605858-41
XX

PD 29-FEB-1996.
XX

XX	16-AUG-1995;	93WU-US010661-
PF		

XX	2408-00258751
EX	25-AUG-1994

EA (UNIM) UNIV WASHINGTON.
PA (STRD) UNIV LELAND STANFORD JUNIOR

06/11/08

XX St Geme JW, Falkow S;
 PI LKGDITQTKGLPFSGRPPTPHAVNHLNKRMEGIPQGLIWDHDMINRTKAFENFOIK 660
 DR WFI, 1996-151147/15.
 DR N-PSDB; AAT17215.
 XX
 PT Haemophilus adhesion and penetration protein and corresponding DNA - used
 XX to produce vaccines against H. influenzae infection.
 PS Claim 2, Fig 6, 105pp; English.
 XX
 CC The sequence represents a Haemophilus influenzae adhesion and penetration
 CC protein, encoded by the hap gene. The protein is first synthesised as a
 CC preprotein with 3 functional domains (the N-terminal signal peptide, the
 CC protease and a C-terminal helper domain), which is transported to the
 CC periplasm, followed by insertion of the C-terminal beta-domain into the
 CC outer membrane, possibly forming a pore, and export of the N-terminal
 CC fragment through the outer membrane, followed by autoproteolytic cleavage
 CC and secretion of the mature protease, leaving an outer membrane protein
 CC fragment. An active site serine residue characteristic of serine
 CC proteases is present, along with 2 cysteine residues which may be
 CC important in forming the catalytic domain. A gene encoding the protein
 CC may be inserted in a vector and expressed in recombinant host cells, for
 CC use as a recombinant vaccine. The protein may also be used in diagnostic
 CC monoclonal antibody production.
 XX
 SC Sequence 1394 AA;
 Query Match 100.0%; Score 7273; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKTVPRHPIKACISLIGISCMAGHTYFGIDYVDPKAKKFTVGAQNTKYNNK 60
 DB 1 MKTVPRHPIKACISLIGISCMAGHTYFGIDYVDPKAKKFTVGAQNTKYNNK 60
 QY 61 GQIVGSMTKAPMIDFVSVNNGVAAVVENQYIVSAHNGYTDVDFGAEQNNPDQRF 120
 DB 61 GQIVGSMTKAPMIDFVSVNNGVAAVVENQYIVSAHNGYTDVDFGAEQNNPDQRF 120
 QY 121 YKIVKNNYKONLHPYEDDYHNRILKFTVTAPEIDMTSNNGSTYSRTIKYPRVRIG 180
 DB 121 YKIVKNNYKONLHPYEDDYHNRILKFTVTAPEIDMTSNNGSTYSRTIKYPRVRIG 180
 QY 181 SGRQFMDNDGDDVAGAVHYLTAGNTNNGAGNGSYLGGVAKAEVPEPLTAGSK 240
 DB 181 SGRQFMDNDGDDVAGAVHYLTAGNTNNGAGNGSYLGGVAKAEVPEPLTAGSK 240
 QY 241 GDSGSPMFIYDAKQKMLINGILREGNPFEGKNGFOLVRKSYFDEIFERDLHTSLYTRA 300
 DB 241 GDSGSPMFIYDAKQKMLINGILREGNPFEGKNGFOLVRKSYFDEIFERDLHTSLYTRA 300
 QY 301 GNGYTTISGNDNGQSITQKGIPESEIKTLANMSLPLKEDKYNRPYDQENIYSPRLN 360
 DB 301 GNGYTTISGNDNGQSITQKGIPESEIKTLANMSLPLKEDKYNRPYDQENIYSPRLN 360
 QY 361 NGETLYFMDQKQSLIFASDINOGAGLIFEGNFTVSNNSQYVQAGHNSNSYTYMK 420
 DB 361 NGETLYFMDQKQSLIFASDINOGAGLIFEGNFTVSNNSQYVQAGHNSNSYTYMK 420
 QY 421 VNGVEHRLSKIGKGLHVAOKGENKGSISVDGKVLLEQADQGNKQAFSELGLVSGR 480
 DB 421 VNGVEHRLSKIGKGLHVAOKGENKGSISVDGKVLLEQADQGNKQAFSELGLVSGR 480
 QY 481 GTVOLANDKQFDTKYFGRGRLDLNHSILTFKRIQNTDYGAMVNHNTQAAVYIT 540
 DB 481 GTVOLANDKQFDTKYFGRGRLDLNHSILTFKRIQNTDYGAMVNHNTQAAVYIT 540
 QY 541 GNESTIVPBGNNINLKYKXIAAYNGFGBTDKNKNGRLALIKPTPDRITLLISGNT 600
 DB 541 GNESTIVPBGNNINLKYKXIAAYNGFGBTDKNKNGRLALIKPTPDRITLLISGNT 600
 QY 601 LKGDITQTKGLPFSGRPPTPHAVNHLNKRMEGIPQGLIWDHDMINRTKAFENFOIK 660

DB 601 LKGDITQTKGLPFSGRPPTPHAVNHLNKRMEGIPQGLIWDHDMINRTKAFENFOIK 660
 QY 661 GGSAAVSPNVSIEGWYTSNNAATGTPVPPNOQNTICRSPMTGLITTCOKVDLDTKYI 720
 DB 661 GGSAAVSPNVSIEGWYTSNNAATGTPVPPNOQNTICRSPMTGLITTCOKVDLDTKYI 720
 QY 721 NSIPKQINSINLTQNAATNKGAKLNGWTLTHSQFTLSNNAQIGNIPLSNNTA 780
 DB 721 NSIPKQINSINLTQNAATNKGAKLNGWTLTHSQFTLSNNAQIGNIPLSNNTA 780
 QY 781 TVDNANLNGVNLTDNAQSLNHSFHOIQDGGTTLLENATWTPSDTLQMLTLNN 840
 DB 781 TVDNANLNGVNLTDNAQSLNHSFHOIQDGGTTLLENATWTPSDTLQMLTLNN 840
 QY 841 STTLNAYSASSNTPRRPSLETFTTSAHRENTLTVNGKLSGGSTQFTSSIFGK 900
 DB 841 STTLNAYSASSNTPRRPSLETFTTSAHRENTLTVNGKLSGGSTQFTSSIFGK 900
 QY 901 SDKLSNDAGDYILSVNTGKEPTELQGLTVESKNDQPLSDKLPRTLENHVDAGL 960
 DB 901 SDKLSNDAGDYILSVNTGKEPTELQGLTVESKNDQPLSDKLPRTLENHVDAGL 960
 QY 961 RYLVKNDGEFRLNPIKQELANDLVAAEQARTLEAKQVPTAKTQTEBPYRSRRA 1020
 DB 961 RYLVKNDGEFRLNPIKQELANDLVAAEQARTLEAKQVPTAKTQTEBPYRSRRA 1020
 QY 1021 RAAPFTLPPDSLLNLEKQAELEAFQSKAKTKKQSKRAVSPDLDOSTPTEEA 1080
 DB 1021 RAAPFTLPPDSLLNLEKQAELEAFQSKAKTKKQSKRAVSPDLDOSTPTEEA 1080
 QY 1081 LEVYDAPQSEKBLAQAEEAKQKQKOLISRSNSALSELATVNSVLSVQDELRLTV 1140
 DB 1081 LEVYDAPQSEKBLAQAEEAKQKQKOLISRSNSALSELATVNSVLSVQDELRLTV 1140
 QY 1141 DQASAVNTNIAQDKRRYDSDA FRAYQOQKTNLROIGVAKALANGRI GAVFSRSNDTE 1200
 DB 1141 DQASAVNTNIAQDKRRYDSDA FRAYQOQKTNLROIGVAKALANGRI GAVFSRSNDTE 1200
 QY 1201 DEGVNHNATLTMGSGPQYQWGDLOFGNVGTCISASVAAEFSQKTHRAIYNGNAY 1260
 DB 1201 DEGVNHNATLTMGSGPQYQWGDLOFGNVGTCISASVAAEFSQKTHRAIYNGNAY 1260
 QY 1261 QFTLGGLGIPYFGVNRPIERENYQSEHRYVTPSLANRNAGIRYDTFTPTNISV 1320
 DB 1261 QFTLGGLGIPYFGVNRPIERENYQSEHRYVTPSLANRNAGIRYDTFTPTNISV 1320
 QY 1321 KPYFPVNVVDSNNAVQTTNLTVALQPFGRYQKEVGLKAEILHFOISAFISKQSQ 1380
 DB 1321 KPYFPVNVVDSNNAVQTTNLTVALQPFGRYQKEVGLKAEILHFOISAFISKQSQ 1380
 QY 1381 GKQNNYGVKLGYN 1394
 DB 1381 GKQNNYGVKLGYN 1394
 RESULT 2
 ADA09343
 ID ADA09343 standard; protein; 1394 AA.
 XX
 AC ADA09343;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Haemophilus adherence and penetration protein (HAP).
 XX
 KW Haemophilus adherence and penetration protein; HAP; antibacterial;
 XX
 KW immunostimulant; vaccine; Haemophilus influenza infection; influenza.
 XX
 OS Haemophilus influenzae.
 XX
 PN US2003073166-A1.
 XX

PD 17-APR-2003.
 XX
 PF 22-FEB-2002; 2002US-00080505.
 XX
 PR 25-AUG-1994; 94US-00296791.
 XX 20-APR-2001; 2001US-0083996.
 XX
 PA (GEME/) GEME J W S.
 PI Gene JMS;
 XX WPI: 2003-567308/53.
 DR N-PSDB; ADA09342.
 XX
 PT New recombinant Haemophilus adhesion and penetration protein, useful as a
 PT vaccine for prophylactic or therapeutic treatment of a Haemophilus
 PT influenzae infections, in generating antibodies for passive immunization.
 PS
 PS Example 1, Fig 6; 149pp; English.
 CC The invention relates to a new recombinant Haemophilus adhesion and
 CC penetration protein (HAP) appearing as ADA09348, ADA09350, ADA09352 and
 CC ADA09354 encoded by a nucleic acid, appearing as ADA09349, ADA09351,
 CC ADA09353, ADA09355 and ADA09357. Also included are a recombinant nucleic
 CC acid that will hybridize under stringency conditions to the nucleic acid
 CC detailed above, an expression vector comprising transcriptional and
 CC translation regulatory nucleic acid operably linked to the HAP nucleic
 CC acid, producing HAP by culturing a host cell transformed with the
 CC expression vector, a composition comprising a pharmaceutical carrier and
 CC HAP, inducing an immune response in a patient by administering the
 CC composition, a composition comprising an antibody capable of binding to
 CC HAP (or to a peptide selected from 35 peptide sequences appearing as
 CC ADA09359-ADA09393), and an antibody capable of binding to the peptides.
 CC The Haemophilus adhesion and penetration (HAP) protein is useful as a
 CC vaccine for prophylactic or therapeutic treatment of an H. influenzae
 CC infection (e.g. influenza), in generating an immune response in a
 CC patient, and in generating antibodies to the HAP protein. The antibodies
 CC may be used to diagnose the presence of H. influenzae infection in a
 CC sample or patient, in passive immunisation treatments, and to purify or
 CC separate HAP proteins or the H. influenzae organism from a sample. The
 CC present sequence is the H. influenzae HAP protein.
 XX
 SQ Sequence 1394 AA;
 Query Match 100.0%; Score 7273; DB 6; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 361 NGETLYFMDOKQSGSLIFASDINQAGGLYFEGNFYSPRSNQTWQAGIHVSENSTVTK 420
 QY 421 VNGVEKHDELSKI GKGLTHVQAKGENKSGISVGGKAVILEQQADDOGNKQAFSEIGVSGR 480
 DB 421 VNGVEHDLRLSKIGKGLTHVQAKGENKSGISVGGKAVILEQQADDOGNKQAFSEIGVSGR 480
 QY 481 GTVQANDKQFDYDKFYFGRGRLDLNGSLTPKRLQNTDEGAVYNNNTQAAVNTT 543
 DB 481 GTVQANDKQFDYDKFYFGRGRLDLNGSLTPKRLQNTDEGAVYNNNTQAAVNTT 540
 QY 541 GNESTVLPNGNINLKDYEKELAVNGAFGETDKKNGKLNLIYKPTTEORTLLSGGN 600
 DB 541 GNESTVLPNGNINLKDYEKELAVNGAFGETDKKNGKLNLIYKPTTEORTLLSGGN 600
 QY 601 LKGDITOTKGLFFSGRPPTPHAYNHLNKRSEMEGIPQELVMDHWINTPFAENFOIK 660
 DB 601 LKGDITOTKGLFFSGRPPTPHAYNHLNKRSEMEGIPQELVMDHWINTPFAENFOIK 660
 QY 661 GSAVAVSRNVSILBGWTVSNNAATFGVVPNOQNTICTRSDMTGLTTCOKVDLTKYI 720
 DB 661 GSAVAVSRNVSILBGWTVSNNAATFGVVPNOQNTICTRSDMTGLTTCOKVDLTKYI 720
 QY 721 NSIPRTQINGSLNTDNATANKGIAKINGVTLTNHSGFTLSNNAIQIGIRLSDNSTA 780
 DB 721 NSIPRTQINGSLNTDNATANKGIAKINGVTLTNHSGFTLSNNAIQIGIRLSDNSTA 780
 QY 781 TVDNANLNGVNLTDLSAOFSLNHSRHOIQGKTTVTLBNATVMSDITLQNTLNN 840
 DB 781 TVDNANLNGVNLTDLSAOFSLNHSRHOIQGKTTVTLBNATVMSDITLQNTLNN 840
 QY 841 STTLNLSAVSASNTPTRRRSLETTTTSAREPNTLVNKLSGGQTFQFSSILFGYK 900
 DB 841 STTLNLSAVSASNTPTRRRSLETTTTSAREPNTLVNKLSGGQTFQFSSILFGYK 900
 QY 901 SDKLKLSNDAGDYLLSYNTGKEPETLEQLTVESKONQPSDKLKTLENDHVTAGAL 960
 DB 901 SDKLKLSNDAGDYLLSYNTGKEPETLEQLTVESKONQPSDKLKTLENDHVTAGAL 960
 QY 961 RYVLVNDGEERLNPRIKQELHNDLVABQAEETLBAQVEPTAKTGYGPKYRSRPA 1020
 DB 961 RYVLVNDGEERLNPRIKQELHNDLVABQAEETLBAQVEPTAKTGYGPKYRSRPA 1020
 QY 1021 RAAPPTLPDQSLNALBAQAEITAFPOKSKATKYVSKRAVSDPLDQSLFALEEA 1080
 DB 1021 RAAPPTLPDQSLNALBAQAEITAFPOKSKATKYVSKRAVSDPLDQSLFALEEA 1080
 QY 1081 LEVTDAPQOESKRLAQEBAEKQKQKDLISRYNSALSATVNSMLSYQDELRLFY 1140
 DB 1081 LEVTDAPQOESKRLAQEBAEKQKQKDLISRYNSALSATVNSMLSYQDELRLFY 1140
 QY 1141 DQASAVWNTLADOKRRYDSDAFRAYVQOKTNLROIGVOKALANGRIGAVFSHSRDNTE 1200
 DB 1141 DQASAVWNTLADOKRRYDSDAFRAYVQOKTNLROIGVOKALANGRIGAVFSHSRDNTE 1200
 QY 1201 DBOYKKAHLLTMGCPQOYQKQDLOFGVNVGTGISASKRAEVSQSKIRKAINVGNAS 1260
 DB 1201 DBOYKKAHLLTMGCPQOYQKQDLOFGVNVGTGISASKRAEVSQSKIRKAINVGNAS 1260
 QY 1261 QFRLGQGLQIPFGVNYFTERENYQSEBRYKPSLAENRYAGIRVDYFTFTDNI 1320
 DB 1261 QFRLGQGLQIPFGVNYFTERENYQSEBRYKPSLAENRYAGIRVDYFTFTDNI 1320
 QY 1321 KPYEFNYVDVSNNAVQTVNLTLVQCPGRYMQVEGLKAEILHIFQISAFISQSGSL 1380
 DB 1321 KPYEFNYVDVSNNAVQTVNLTLVQCPGRYMQVEGLKAEILHIFQISAFISQSGSL 1380
 QY 1381 GKQQNTGYCLGTRYK 1394
 DB 1381 GKQQNTGYCLGTRYK 1394
 QY 361 NGETLYFMDOKQSGSLIFASDINQAGGLYFEGNFYSPRSNQTWQAGIHVSENSTVTK 420

RESULT 3

ADAO9348
 ID ADAO9348 standard; protein, 1395 AA.
 AC ADAO9348;
 DT 06-NOV-2003 (first entry)
 DE Haemophilus HAP protein from strain N187.
 XX Haemophilus adherence and penetration protein; HAP; antibacterial;
 KW immunostimulant; vaccine; Haemophilus influenza infection; influenza.
 XX Haemophilus influenzae; strain N187.
 CS Haemophilus influenzae; strain N187.
 PN US2003073166-A1.
 PD 17-APR-2003.
 XX 22-FEB-2002; 2002US-00080505.
 XX 25-AUG-1994; 94US-00296791.
 PR 20-APR-2001; 2001US-00839996.
 PA (GENE) / GENE J W S.
 PI Gene JWS;
 DR MPI; 2003-567308/53.
 XX The invention relates to a new recombinant Haemophilus adhesion and
 PT vaccine for prophylactic or therapeutic treatment of a Haemophilus
 PT influenzae infection, in generating antibodies for passive immunization.
 PS Claim 2; Fig 11; 14pp; English.
 CC The invention relates to a new recombinant Haemophilus adhesion and
 CC penetration protein (HAP) appearing as ADAO9348, ADAO9350, ADAO9352 and
 CC ADAO9354 encoded by a nucleic acid appearing as ADAO9349, ADAO9351,
 CC ADAO9353, ADAO9355 and ADAO9357. Also included are a recombinant nucleic
 CC acid that will hybridize under stringency conditions to the nucleic acid
 CC detailed above, an expression vector comprising transcriptional and
 CC translation regulatory nucleic acid operably linked to the HAP nucleic
 CC acid, producing HAP by culturing a host cell transformed with the
 CC expression vector, a composition comprising a pharmaceutical carrier and
 CC HAP, inducing an immune response in a patient by administering the
 CC composition, a composition comprising an antibody capable of binding to
 CC HAP (or to a peptide selected from 35 peptide sequences appearing as
 CC ADAO9353-ADAO9357), and an antibody capable of binding to the peptides.
 CC The Haemophilus adhesion and penetration (HAP) protein is useful as a
 CC vaccine for prophylactic or therapeutic treatment of an H. influenzae
 CC infection (e.g. influenza), in generating an immune response in a
 CC patient, and in generating antibodies to the HAP protein. The antibodies
 CC may be used to diagnose the presence of H. influenzae infection in a
 CC sample or patient, in passive immunisation treatments, and to purify or
 CC separate HAP proteins or the H. influenzae organism from a sample. The
 CC present sequence is the HAP protein from a strain of H. influenzae.
 XX Sequence 1395 AA;
 SO
 Query Match 99.6%; Score 7246.5; DB 6; Length 1395;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1393; Conservative 0; Mismatches 0; Indels 3; Gaps 2;
 QY 1 MKKTVPFLNLFACISLIGISQAMAGTFFGIDYQYRFAENKGFYGAQNIKYNKQ 60
 DB 1 MKKTVPFLNLFACISLIGISQAMAGTFFGIDYQYRFAENKGFYGAQNIKYNKQ 60
 QY 61 GQIVGTSMTKAPMIDSVVSHNGVAALVENVQYIVSAHNGYTDVFGAEGNPDQRF 120
 DB 61 GQIVGTSMTKAPMIDSVVSHNGVAALVENVQYIVSAHNGYTDVFGAEGNPDQRF 120
 QY 121 YKIVKNNYKKNLHPEYDDYHNPRLHKEVTEAPIDMTSNMNGSTYSRTKYPERRVIG 180
 DB 121 YKIVKNNYKKNLHPEYDDYHNPRLHKEVTEAPIDMTSNMNGSTYSRTKYPERRVIG 180

DB 121 YKIVKNNYKKNLHPEYDDYHNPRLHKEVTEAPIDMTSNMNGSTYSRTKYPERRVIG 180
 QY 181 SGRQFMNDODKGDYVAGAYHYLTAGTTHNORGANGSYVZGQVRRAGEYGLPZAGSK 240
 DB 181 SGRQFMNDODKGDYVAGAYHYLTAGTTHNORGANGSYVZGQVRRAGEYGLPZAGSK 240
 QY 241 GDSGPMFTYDAKQKWLINGILBGNPFEGKENGQVLRKSYFDEIFERDHTSLYTRA 300
 DB 241 GDSGPMFTYDAKQKWLINGILBGNPFEGKENGQVLRKSYFDEIFERDHTSLYTRA 300
 QY 301 GNGVYTTSGNDNCGSITQKSGIPEIKITLAWMSLPLEKDXKHNRYPYSPRLN 360
 DB 301 GNGVYTTSGNDNCGSITQKSGIPEIKITLAWMSLPLEKDXKHNRYPYSPRLN 360
 QY 361 NGETLYFMDQKGSILPASDINOAGAGLYFEGNFTVSPNSNQWQAGIHVSESTVTWK 420
 DB 361 NGETLYFMDQKGSILPASDINOAGAGLYFEGNFTVSPNSNQWQAGIHVSESTVTWK 420
 QY 421 VNGVEHDLSTKIGKTLAVQAKGENKGSISVPGKTYLEQADQGNKQAFSEIGVSGR 480
 DB 421 VNGVEHDLSTKIGKTLAVQAKGENKGSISVPGKTYLEQADQGNKQAFSEIGVSGR 480
 QY 481 GTVQIANDKQPDMDKYPFSGRGLLNGHSLFPRKIQNTDGBAMIVNHTTOAANVTIT 540
 DB 481 GTVQIANDKQPDMDKYPFSGRGLLNGHSLFPRKIQNTDGBAMIVNHTTOAANVTIT 540
 QY 541 GNESTVLPNGNNINKLDYRKEIAYNGWFGETDKKNGRLNLKYPTEEDRLILLGGIN 600
 DB 541 GNESTVLPNGNNINKLDYRKEIAYNGWFGETDKKNGRLNLKYPTEEDRLILLGGIN 600
 QY 601 LKGDITQKGLTFSGRPPTPHAYNHLKRSBMEGIPQSELYWHDVINTPEAEVQIK 660
 DB 601 LKGDITQKGLTFSGRPPTPHAYNHLKRSBMEGIPQSELYWHDVINTPEAEVQIK 660
 QY 661 GGSAAVSRVSSIEGMMTVSNNAATRGVPMQNTICRSPMGLTQCKVLLDTTKYI 720
 DB 661 GGSAAVSRVSSIEGMMTVSNNAATRGVPMQNTICRSPMGLTQCKVLLDTTKYI 720
 QY 721 NSIPYQZNSGILTNATANYKGLAKLNGVTLTNHSGFTLSNNATQIGNIRLSDNSTA 780
 DB 721 NSIPYQZNSGILTNATANYKGLAKLNGVTLTNHSGFTLSNNATQIGNIRLSDNSTA 780
 QY 781 TYDNANLNGVHLLTDSAGSLNSHSHQIQGKGTVTLNMAITWPSOTITQNTLNN 840
 DB 781 TYDNANLNGVHLLTDSAGSLNSHSHQIQGKGTVTLNMAITWPSOTITQNTLNN 840
 QY 841 STITLNSAYSSASNNP--RRSLETTPTSAEHRFNTLVNGKLSQGTPOFTSILFG 898
 DB 841 STITLNSAYSSASNNP--RRSLETTPTSAEHRFNTLVNGKLSQGTPOFTSILFG 898
 QY 899 YKSDKLKLSNDASQVYLSVNTGKEPETLEOUTLVESKXNPISDKKFTLENDHWAG 958
 DB 899 YKSDKLKLSNDASQVYLSVNTGKEPETLEOUTLVESKXNPISDKKFTLENDHWAG 958
 QY 901 YKSDKLKLSNDASQVYLSVNTGKEPETLEOUTLVESKXNPISDKKFTLENDHWAG 960
 DB 901 YKSDKLKLSNDASQVYLSVNTGKEPETLEOUTLVESKXNPISDKKFTLENDHWAG 960
 QY 959 ALRYKLVDNDEPRLNPTKEQELANDLVAREQAEPTLAKQVEPTAKTGTGEPKRSR 1018
 DB 959 ALRYKLVDNDEPRLNPTKEQELANDLVAREQAEPTLAKQVEPTAKTGTGEPKRSR 1018
 QY 961 ALRYKLVDNDEPRLNPTKEQELANDLVAREQAEPTLAKQVEPTAKTGTGEPKRSR 1020
 DB 961 ALRYKLVDNDEPRLNPTKEQELANDLVAREQAEPTLAKQVEPTAKTGTGEPKRSR 1020
 QY 1019 AALEVIDAPQSEKRLAGEAEKQKQKOLISRYNSNLSLSLSTVNSMLSYQDELRL 1138
 DB 1019 AALEVIDAPQSEKRLAGEAEKQKQKOLISRYNSNLSLSLSTVNSMLSYQDELRL 1138
 QY 1139 FVDQNSAWNTNIPQKREYSDAPRAVOQKTMISQICWYLLANRISAVSSESSEN 1198
 DB 1139 FVDQNSAWNTNIPQKREYSDAPRAVOQKTMISQICWYLLANRISAVSSESSEN 1198
 QY 1141 FVDQNSAWNTNIPQKREYSDAPRAVOQKTMISQICWYLLANRISAVSSESSEN 1199
 DB 1141 FVDQNSAWNTNIPQKREYSDAPRAVOQKTMISQICWYLLANRISAVSSESSEN 1199
 QY 1199 TPEQVQKHAULTMWSGFAQYQMGQJCFGVVGTGTSASKMAEQQSKIRKAINGVNA 1258
 DB 1199 TPEQVQKHAULTMWSGFAQYQMGQJCFGVVGTGTSASKMAEQQSKIRKAINGVNA 1258
 QY 1200 TPEQVQKHAULTMWSGFAQYQMGQJCFGVVGTGTSASKMAEQQSKIRKAINGVNA 1259
 DB 1200 TPEQVQKHAULTMWSGFAQYQMGQJCFGVVGTGTSASKMAEQQSKIRKAINGVNA 1259

QY 1259 SYGRLGQLGLOPYFYGVNRYFIRENVOSEEVKVTSLAFNRNAGIRVDYETPTDNT 1318
 DB 1260 SYGRLGQLGLOPYFYGVNRYFIRENVOSEEVKVTSLAFNRNAGIRVDYETPTDNT 1319
 QY 1319 SYVPYFVNVVDSNANVQTTNLTVLQOPFGRYMOKEVGLAKELHFOISAFISQSS 1378
 DB 1320 SYVPYFVNVVDSNANVQTTNLTVLQOPFGRYMOKEVGLAKELHFOISAFISQSS 1379
 QY 1379 QLGKQONVGVKLGTRM 1394
 DB 1380 QLGKQONVGVKLGTRM 1395
 RESULT 4
 ADA09350
 ID ADA09350 standard; protein; 1434 AA.
 XX ADA09350;
 AC
 XX
 XX 06-NOV-2003 (first entry)
 DT
 XX
 DE Haemophilus HAP protein from strain 11.
 XX
 XX Haemophilus adherence and penetration protein; HAP; antibacterial;
 KM immunostimulant; vaccine; Haemophilus influenza infection; influenza.
 XX
 OS Haemophilus influenzae; strain 11.
 XX
 PN US2003073166-A-
 XX
 PD 17-APR-2003.
 XX
 XX 22-FEB-2002; 2002US-00080505.
 XX
 PR 25-AUG-1994; 94US-00296791.
 PR 20-APR-2001; 2001US-0083996.
 XX
 XX (GEMF/) GEME J W S.
 PA
 XX
 PI GEME JMS;
 DR WPI; 2003-567308/53.
 DR N-PSDB; ADA09349.
 XX
 PT New recombinant Haemophilus adhesion and penetration protein, useful as a
 PT vaccine for prophylactic or therapeutic treatment of a Haemophilus
 PT influenzae infections, in generating antibodies for passive immunization.
 XX
 PS Claim 2; Fig 17; 149pp; English.
 XX
 CC The invention relates to a new recombinant Haemophilus adhesion and
 CC penetration protein (HAP) appearing as ADA09348, ADA09350, ADA09352 and
 CC ADA09354 encoded by a nucleic acid appearing as ADA09349, ADA09351,
 CC ADA09353, ADA09355 and ADA09357. Also included are a recombinant nucleic
 CC acid that will hybridize under stringency conditions to the nucleic acid
 CC detailed above, an expression vector comprising transcriptional and
 CC translation regulatory nucleic acid operably linked to the HAP nucleic
 CC acid, producing HAP by culturing a host cell transformed with the
 CC expression vector, a composition comprising a pharmaceutical carrier and
 CC HAP, inducing an immune response in a patient by administering the
 CC composition, a composition comprising an antibody capable of binding to
 CC HAP (or to a peptide selected from 35 peptide sequences appearing as
 CC ADA09359-ADA09363), and an antibody capable of binding to the peptides.
 CC The Haemophilus adhesion and penetration (HAP) protein is useful as a
 CC vaccine for prophylactic or therapeutic treatment of an H. influenzae
 CC infection (e.g. influenza), in generating an immune response in a
 CC patient, and in generating antibodies to the HAP protein. The antibodies
 CC may be used to diagnose the presence of H. influenzae infection in a
 CC sample or patient, in passive immunisation treatments, and to purify or
 CC separate HAP proteins or the H. influenzae organism from a sample. The
 CC present sequence is the HAP protein from a strain of H. influenzae.

SQ Sequence 1434 AA;
 Query Match 84.1%; Score 6115; DB 6; Length 1434;
 Best Local Similarity 82.9%; Pred. No. 0;
 Matches 1194; Conservative 81; Mismatches 113; Indels 52; Gaps 10;
 QY 1 MCKTVFRLNLTACISLGVISGMAAGTYYGIDYOTRDPAENKGTTCAGQNIKVTYKQ 60
 DB 1 MCKTVFRLNLTACISLGVISGMAAGTYYGIDYOTRDPAENKGTTCAGQNIKVTYK 60
 QY 61 GOLVTSMTAKPMIDSVSVSRNGVAALVENQYIVSAHVAVGTVDGARGNPDQHF 120
 DB 61 GOLVTSMTAKPMIDSVSVSRNGVAALVENQYIVSAHVAVGTVDGARGNPDQHF 120
 QY 121 YKIVKNNYKKDNLAPYEDDYNNPRLHKFTFAPIIDMTSNMNGSTYSRRTKYPERV 180
 DB 121 YKIVKNNYKKDNLAPYEDDYNNPRLHKFTFAPIIDMTSNMNGSTYSRRTKYPERV 180
 QY 181 SGRQFWRNDQKGDQVAGAYHYLTNGTNNQAGAGISYLAGDYRACAGYGLPTASSK 240
 DB 181 SGRQFWRNDQKGDQVAGAYHYLTNGTNNQAGAGISYLAGDYRACAGYGLPTASSK 240
 QY 241 GDSGPMFTIDAEQMLNGLVLRBQPYGDFGQLAKSFLDEIRKDGNGFLTPK 300
 DB 241 GDSGPMFTIDAEQMLNGLVLRBQPYGDFGQLAKSFLDEIRKDGNGFLTPK 300
 QY 301 GNGVYTTISGNDGQSGITQSGIPSEIKITLANNMSPLEKDXKHPRYDGPNTYSPRLN 360
 DB 301 GNGVYTTISGNDGQSGITQSGIPSEIKITLANNMSPLEKDXKHPRYDGPNTYSPRLN 360
 QY 361 NSELTLYEMDQKQSLIPASDINQAGGLYFENFTYSPNSQNTQAGGIVSNSTVYWK 420
 DB 361 NSELTLYEMDQKQSLIPASDINQAGGLYFENFTYSPNSQNTQAGGIVSNSTVYWK 420
 QY 421 VNGVENDELSTKIGKGLTHQAKGKNSISVSDKYLIEQQAQDQKQKQAFSELTGVSR 480
 DB 421 VNGVENDELSTKIGKGLTHQAKGKNSISVSDKYLIEQQAQDQKQKQAFSELTGVSR 480
 QY 481 GTVOLNDKQPDPTDKYFPGFRGRJLDNGHSLTFKXIONTEGAMTVNHTTQAAVNT 540
 DB 481 GTVOLNDKQPDPTDKYFPGFRGRJLDNGHSLTFKXIONTEGAMTVNHTTQAAVNT 540
 QY 541 GNESTVLP-NGNNINKLYREKELIANGNFGEDTQNGKRLNLYKPTEDRLLTSGGT 599
 DB 541 GNESTVLP-NGNNINKLYREKELIANGNFGEDTQNGKRLNLYKPTEDRLLTSGGT 599
 QY 599 NIKGPIITQKGLFPGGRPTPHAVVHLNKRSEMEGIPQGEIYMDHPNRTFAENFQI 659
 DB 599 NIKGPIITQKGLFPGGRPTPHAVVHLNKRSEMEGIPQGEIYMDHPNRTFAENFQI 659
 QY 660 KGSAAVSRNYSIENGTWSNNANATPGVVPQONTICIRSDWTGLTTCQKYLDTTRY 719
 DB 660 KGSAAVSRNYSIENGTWSNNANATPGVVPQONTICIRSDWTGLTTCQKYLDTTRY 719
 QY 719 INSIEKTQINGSINLTDNATANKGLAVKAGNTVLTNHSQFTLSNATQIGNRIS- 775
 DB 719 INSIEKTQINGSINLTDNATANKGLAVKAGNTVLTNHSQFTLSNATQIGNRIS- 775
 QY 776 ---DNST-----ATVDNANGVHLTDSAP 800
 DB 776 ---DNST-----ATVDNANGVHLTDSAP 800
 QY 801 LKNSHSHQIGDKGTTVTLENATWTPSDTTLONTLANSITLANSYASASSNTP-RR 859
 DB 801 LKNSHSHQIGDKGTTVTLENATWTPSDTTLONTLANSITLANSYASASSNTP-RR 859
 QY 859 LKNSHSHQIGDKGTTVTLENATWTPSDTTLONTLANSITLANSYASASSNTP-RR 898
 DB 859 LKNSHSHQIGDKGTTVTLENATWTPSDTTLONTLANSITLANSYASASSNTP-RR 898
 QY 898 RSLIETITPTSAERFNTLVNGKLSGGTQFPSPISFGEISDKKLSNDAEDGYTLTVR 958
 DB 898 RSLIETITPTSAERFNTLVNGKLSGGTQFPSPISFGEISDKKLSNDAEDGYTLTVR 958
 QY 920 NCKEETTELQTLVSKONQPLSDKLFTELNDHDAAGLKYKLVKIDGFRLLNPDI 979
 DB 920 NCKEETTELQTLVSKONQPLSDKLFTELNDHDAAGLKYKLVKIDGFRLLNPDI 979
 QY 959 NTGKEFTVLEQLTVESKDKPLSDKLFTELNDHDAAGLKYKLVKIDGFRLLNPDI 1018
 DB 959 NTGKEFTVLEQLTVESKDKPLSDKLFTELNDHDAAGLKYKLVKIDGFRLLNPDI 1018


```

QY 781 TVDNANLNGNVLHLDSDAQLKNSHSHQIQDGKTVTLLENKVTMTESDTLLQNTLNN 840
DB 773 TVDNANLNGNVLHLDSDAQLKNSHSHQIQDGKTVTLLENKVTMTESDTLLQNTLNN 832
QY 841 STTLNAYSASSNNTR-RRSLTETTTYSNERETTLTVNGKLSGGGTQFTSSLSFG 899
DB 833 STVTLNAYSASSNNAPRRHRSLETETTTYSNERETTLTVNGKLSGGGTQFTSSLSFG 892
QY 900 KSKKTLSDNDAGDYTLISVNTGKPELTLEQTLVYESKQPLSDKLFLENDVNDAGA 959
DB 893 KSKKTLSDNDAGDYTLISVNTGKPELTLEQTLVYESKQPLSDKLFLENDVNDAGA 952
QY 960 LRYKLVNDGFEFLNPLTXEQELNDLVYAEQARTLEAKQVEPTAKTYGEPVRSRA 1019
DB 953 WRKELVKNGEFLNPLTXEQELNDLVYAEQARTLEAKQVEPTAKTYGEPVRSRA 1010
QY 1020 ARAAPDTLPDQSLNLEAKQELTAEQVSKA-----KTKVRSK---AV 1064
DB 1011 -RAAFDTPDQSLNLEAKQELTAEQVSKA-----KTKVRSK---AV 1069
QY 1065 FSDPLDQSL-----FA-----LEALVETDAPQOSKRLAQEERQKQKQ 1108
DB 1070 LATAVEKETTAQIDPANAQAQLNLTQELERLAV---AQAQEKERKAQEQAKQKQKQ 1126
QY 1109 LISRYNSALSELATVNSMLSVQDELDELFDQAQSAVNTNTAOKRKYDSDAFRAVQ 1168
DB 1127 LISRYNSALSELATVNSMLSVQDELDELFDQAQSAVNTNTAOKRKYDSDAFRAVQ 1185
QY 1169 QKTNLQIQVQKALNGLICAVNHSRSDNTPEQVNHATLTMMSGAPQVOMQDLOFGV 1228
DB 1186 QKTNLQIQVQKALNGLICAVNHSRSDNTPEQVNHATLTMMSGAPQVOMQDLOFGV 1245
QY 1229 NVGTGASAKMAEEOSRKIRKAIKNTGVNASYQFELQQLGQPFYGVNRYELERENTQSE 1288
DB 1246 NVGTGASAKMAEEOSRKIRKAIKNTGVNASYQFELQQLGQPFYGVNRYELERENTQSE 1305
QY 1289 EVRYKTPSLAFNRNAGIRVDYTPPTDNISVYKPYFFVNVYVSNANVGTVNLTVLQOP 1348
DB 1306 EVRYKTPSLAFNRNAGIRVDYTPPTDNISVYKPYFFVNVYVSNANVGTVNLTVLQOP 1365
QY 1349 FGRYWCQEVGLQKELIHFQISAFISKSQSGQSGKQKQNVYGLGYR 1394
DB 1366 FGRYWCQEVGLQKELIHFQISAFISKSQSGQSGKQKQNVYGLGYR 1411

```

RESULT 6
ADA09354 ID ADA09354 standard; protein; 1436 AA.

XX AC ADA09354;
XX DB 06-NOV-2003 (first entry)
XX DE Haemophilus HAP protein from strain 860295.
XX KM Haemophilus adherence and penetration protein; HAP; antibacterial;
XX KW immunostimulant; vaccine; Haemophilus influenza infection; influenza.
XX OS Haemophilus influenzae; strain 860295.
XX PN US2003073166-A1.
XX PD 17-APR-2003.
XX PF 22-FEB-2002; 2002US-00080505.
XX PR 25-AUG-1994; 94US-00296791.
XX PR 20-APR-2001; 2001US-00839936.
XX PA (GENE/) GENE J W S.
XX PI Geme JWS;

```

XX WP1; 2003-567306/53.
DR N-PSDB; ADA09353.
XX New recombinant Haemophilus adhesion and penetration protein, useful as a
PT vaccine for prophylactic or therapeutic treatment of a Haemophilus
PT influenzae infections, in generating antibodies for passive immunization.
PS Claim 2; Fig 21; 149pp; English.
XX The invention relates to a new recombinant Haemophilus adhesion and
XX penetration protein (HAP) appearing as ADA09348, ADA09350, ADA09352 and
XX ADA09354 encoded by a nucleic acid appearing as ADA09349, ADA09351,
XX ADA09353, ADA09355 and ADA09357. Also included are a recombinant nucleic
XX acid that will hybridize under stringency conditions to the nucleic acid
XX detailed above, an expression vector comprising transcriptional and
XX translation regulatory nucleic acid operably linked to the HAP nucleic
XX acid, producing HAP by culturing a host cell transformed with the
XX expression vector, a composition comprising a pharmaceutical carrier and
XX HAP, inducing an immune response in a patient by administering the
XX composition, a composition comprising an antibody capable of binding to
XX HAP (or to a peptide selected from 35 peptide sequences appearing as
XX ADA09359-ADA09393), and an antibody capable of binding to the peptides.
XX The Haemophilus adhesion and penetration (HAP) protein is useful as a
XX vaccine for prophylactic or therapeutic treatment of an H. influenzae
XX infection (e.g. influenza), in generating an immune response in a
XX patient, and in generating antibodies to the HAP protein. The antibodies
XX may be used to diagnose the presence of H. influenzae infection in a
XX sample or patient, in passive immunisation treatments, and to purify or
XX separate HAP proteins or the H. influenzae organism from a sample. The
XX present sequence is the HAP protein from a strain of H. influenzae.
SQ Sequence 1436 AA;
Query Match 79.8%; Score 5807; DB 6; Length 1436;
Beat Local Similarity 79.3%; Pred. No. 0;
Matches 1149; Conservative 78; Mismatches 154; Indels 68; Gaps 14;

```

1 MKKTVFRLNPLTACISLIGVSOAMAGHTYRGIDYRPFANCKKFTVGAQKVIKYNQ 60
1 MKKTVFRLNPLTACISLIGVSOAMAGHTYRGIDYRPFANCKKSVAKRIEYVNE 60

61 GOLNCTSMTKAPMIDSVSRNGVALYENQYIVSAHVGVTDVDCAGANPDQHRT 120
61 GLVNGISMTKAPMIDSVSRNGVALYGVQYIVSAHVGVTDVDCAGANPDQHRT 120

121 YKIVKNNYKKKCNLHPEADYHNPRLNKEVTEAPIDMTSNNGSTYSDRTKYPERVIG 180
121 YKIVKNNYKKKGNPRLHGDYHNPRLHKEVTEADPAKMTDMNNGVADLSKYPDRVRIG 180

181 SGRQFNRNDQ-----KQDVAGAYHYTLAGNTHNQRGAGNGSYLSGQVRAKGEVGLP 235
181 TGEQWFRDDBEQKQSKSFWLADATLWRIAGNTHSOSGAGNCTVNLSDTKNNNGVGLP 240

236 IAGSGDGSQPMFYDAKQKWLINGIIRSGNPEPKENGFLQVAKSYF-DEIFERDLHT 294
241 TGVSFDSQSPFTYDAIKQKWLINGVLOTGNPSPSAGNGFLQIKKMTFNFTVEDLPI 300

295 SLTYRAGNVYTIISGNDNGSGSIYQ---KSGIPEIKITLANSPLLEKQVHNAPRY-- 349
301 TFLFPRSNHGSFTSNNGGTGVTQTNNEKVSMPQFKVRFVQFNELAKQK--ZEVYAA 358

350 DEPNITSPRLNGETLYFNDQKQSLIPASDINGAGGIFYEGNTTNSNNOVQAGI 409
359 GGVNAYKPELNNKNIYFGRGIGTLIENININGAGGIFYEGNTTNSNNOVQAGV 418

410 HVSNGSTVPMKNGVHEKRLSKIGGLTAVQAKENGSGISVGDGVTLLEQADQDQNG 469
419 HVEDSTVPMKNGVHEKRLSKIGGLTAVQAKENGSGISVGDGVTLLEQADQDQNG 478

470 AFSEIGLVSGRTYQVNDNDKQPDTKFYFGRGRLDNGSLTFKRIQNTDSGAMIVNE 529
479 AFKEVGVISGRATYQVNSADQVDPNNIYFGRGRLDNGSLTFKRIQNTDSGAMIVNE 538

```

QY 530 NTCQANVTITGNESTVLP--NGNNINKLDVRKEIAVNGMFGEDTKNKNKGRNLITVPTT 588
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 539 NTVVAVNTITIGNESTVAPSKNNINKLDYSKEIAYKMGFEEDENKNGRNLITVPTT 588
QY 588 EDRTLLASGTLNKGDITOTKGLFPSPGRTPPHAVNLNKRKSMEMGIPQGEIYWDHWT 648
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 589 EDRTLLASGTLNKGNITQEGTLVFGSRTPPHAVNLN--RPNEL--GRPGSVIVDDMI 656
QY 649 NRTKANPQIKGSAVSVSRVSSIEGNVTNSNNANAFPGVPRQONTCTRSMTGLTT 708
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 657 TRTPKANPQIKGSAVSVSRVSSIEGNVTNSNNANAFGVVPQONTCTRSMTGLTT 716
QY 709 CQKVDLPTKVINSIPKQIINGSINTLQNAITANVGLAKLNGVTLTNESQPTLSNNAQ 768
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 717 CKYVDLPTKVINSIPTQINGSINTLQNAITANVHGLAKLNGVTLTNESQPTLSNNAQ 776
QY 769 IGNIRLSDNSTATVQANLNGNVALTDSAQFSIKNSHPSHQIQDGKPTVLENAVW 828
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 777 TGNITQLSNHNATVQANLNGNVALTDSAQFSIKNSHPSHQIQDGKPTVLENAVW 836
QY 829 SDTTLQNLTLNNTITLNSAYSASSNNTTP--RSLPTETPTSHREPTLVNNGKLSGQ 887
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 837 SDATLQNLTLNNTITLNSAYSASSNNAPEHRSLEETPTSHREPTLVNNGKLSGQ 896
QY 888 GTFOFTSLFGYKSDKLKLSNDADGYLISVNTKKEBETLEQTLVESKDNQPLSDKLX 947
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 897 GTFOFTSLFGYKSDKLKLSNDADGYLISVNTKKEBETLEQTLVESKDNQPLSDKLX 956
QY 946 FTLENDVDAALRYLVKNDGEFRLNPIKEQELHNDLVABQAEERTLBAQVEFTKNT 1007
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 957 FTLENDVDAALRYLVKNDGEFRLNPIKEQELHNDLVABQAEERTLEAQVEFTKNT 1016
QY 1008 QNGEPKRSRRARAFPTLPDQSLNLEKQAELEAFQSKATKTKNSKRAVSD 1067
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1017 QNSARVRSK--RAVSDTLPDQSLDVLQAEQVEPFAEKQNKKA--KKVSKRAVSD 1071
QY 1068 PLTDOSS-----LFALEALEVID 1085
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1072 TLTPDQSLDVLQAEQVEPFAEKQNKAKYKRSKRAAFPTLPDLSRLKYLEVLAIVN 1131
QY 1086 APOQSEKURLAOEBAEKQKQKQDLISRYNSALSISLSTVNSMLSVDELDLFPVDAQS 1145
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1132 AQQQYKKE---PODQEKQKQKQDLISRYNSALSISLSTVNSMLSVDELDLFPVDAQS 1188
QY 1146 AAMNTIAQDERRVDSQAFPAVQCCQKTLRQIGVQALANGRIGAVSHSRSDNTPEQVK 1205
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1189 AAMNTIAQDERRVDSQAFPAV--QCKTWRQIGVQALANGRIGAVSHSRSDNTPEQVK 1247
QY 1206 NHATLTMMSGFAQYQWGDLOFGVAVNGTGISASMAEBSRKIHKKAINTGVNASYQFPLG 1265
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1248 NHATLTMMSGFAQYQWGDLOFGVAVNGTGISASMAEBSRKIHKKAINTGVNASYQFPLG 1307
QY 1266 QLGQIPYGVNRYFIRENYSQSEVRYKTSIAFNRYNAGRIVYTPPTDNIISYKPYFE 1325
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1308 QLGQIPYGVNRYFIRENYSQSEVRYKTSIAFNRYNAGRIVYTPPTDNIISYKPYFE 1367
QY 1326 VAVYVVSANAVQTTVAIVLQOPFGRYWQKEVGLAKETLHPISAFISQSOSQULGXON 1385
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1368 VAVYVVSANAVQTTVAIVLQOPFGRYWQKEVGLAKETLHPISAFISQSOSQULGXON 1427
QY 1386 VGVKLGYRM 1394
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1428 VGVKLGYRM 1436

```

RESULT 7
ADA09356
ADA09356 standard; protein, 1391 AA.

XX ADA09356;
DT 06-NOV-2003 (first entry)

```

XX Haemophilus HAP protein from strain 3219B.
DE
XX
KW Haemophilus adherence and penetration protein; HAP; antibacterial;
KW Immunostimulant; vaccine; Haemophilus influenza infection; influenza.
XX
XX Haemophilus influenzae; strain 3219B.
OS
XX
XX US2003073166-A1.
PN
XX
XX 17-Apr-2003.
PD
XX
XX 22-FEB-2002; 2002US-00080505.
PF
XX
XX 25-AUG-1994; 94US-00296791.
PR 20-Apr-2001; 2001US-00839996.
XX
XX (GENE/J) GENE J W S.
PA
XX
PI Gene JWS;
XX
XX MPI: 2003-567308/53.
DR N-PEDB; ADA09355.
XX
XX New recombinant Haemophilus adhesion and penetration protein, useful as a
PT vaccine for prophylactic or therapeutic treatment of a Haemophilus
PT influenzae infections, in generating antibodies for passive immunization.
XX
XX Example 2; Fig 23; 14pp; English.
PS
XX
XX The invention relates to a new recombinant Haemophilus adhesion and
CC penetration protein (HAP) appearing as ADA09348, ADA09350, ADA09352 and
CC ADA09354 encoded by a nucleic acid appearing as ADA09349, ADA09351,
CC ADA09353, ADA09355 and ADA09357. Also included are a recombinant nucleic
CC acid that will hybridize under stringency conditions to the nucleic acid
CC detailed above, an expression vector comprising transcriptional and
CC translation regulatory nucleic acid operably linked to the HAP nucleic
CC acid, producing HAP by culturing a host cell transformed with the
CC expression vector, a composition comprising a pharmaceutical carrier and
CC HAP, inducing an immune response in a patient by administering the
CC composition, a composition comprising an antibody capable of binding to
CC HAP (or to a peptide selected from 35 peptide sequences appearing as
CC ADA09339-ADA09393), and an antibody capable of binding to the peptides.
CC The Haemophilus adhesion and penetration (HAP) protein is useful as a
CC vaccine for prophylactic or therapeutic treatment of an H. influenzae
CC infection (e.g. influenza), in generating an immune response in a
CC patient, and in generating antibodies to the HAP protein. The antibodies
CC may be used to diagnose the presence of H. influenzae infection in a
CC sample or patient, in passive immunisation treatments, and to purify or
CC separate HAP proteins or the H. influenzae organism from a sample. The
CC present sequence is the HAP protein from a strain of H. influenzae.
XX
XX Sequence 1391 AA:
SQ

```

Query Match 72.9%; Score 5303.5; DB 6; Length 1391;
Best Local Similarity 74.8%; Pred. No. 8.1e-297;
Matches 1053; Conservative 109; Mismatches 215; Indels 31; Gaps 15;

```

QY 1 MKKTVFRLNELTACISLGIYSQAMAGTYFGIDYQYRDAENKGFYTGQAKIKYVKRQ 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MKKTVFRLNELTACISLGIYSQAMAGTYFGIDYQYRDAENKGFYTGQAKIKYVKRQ 60
QY 61 GOLVGTSMKAPMIDFSVSRNGVALYENQYIVSAHNVGYTDVDFGAGBNNDQHRFT 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 GEMTGTMMKQVMPDLSVWVGYSITLSSEQHLLSVAHNVGYDVDFGMEGENDQHRFX 120
QY 121 YKIVKKNYKKQNLHPFEDDYANPRLKKTYYTBAFZIMTSNNAAGSYSDRKYPERRATG 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 YKIVKKNYKXSGDRQ--YNDYQHPRLKKTYYTBAFIMVSYDGNHKKKQYELRYRVG 178
QY 181 SGRQFMRNDQK--GDQVAGAYHYTLAGTNRQAGAGNGSYLGGVRAAGGYPLPIAG 238
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 SGRQFMRNDQK--GDQVAGAYHYTLAGTNRQAGAGNGSYLGGVRAAGGYPLPIAG 237

```


QY 1 MKXTFRLNPLACISLGIVSOAMAGHTYFGIDYOYRDFAEKNGKFTYGAQNIKXVNO 60
 DB 1 MKXTFRLNPLACISLGIVSOAMAGHTYFGIDYOYRDFAEKNGKFTYGAQNIKXVNO 60
 QY 61 GOLVTSMTKAMIDFSVSRNGVAALVENOYIVSAHNGYTDVDFAGBNNDQHRPT 120
 DB 61 GEMIGTMKGVMPPLDSMGVRCGYSTLISEGLISVAHNGYDVDFGMEGNPDQHRP 120
 QY 121 YKIVKRNNTKONLHPEDDYHNPRLKFTYTAAPIDMTSNMNGSYSDRTKTYERVRIG 180
 DB 121 YKIVKRNNTKONLHPEDDYHNPRLKFTYTAAPIDMTSNMNGSYSDRTKTYERVRIG 180
 QY 181 SGOQFNRNDQK--GDVAGAHYHLLRAGTNHOGAGNGSYSGYQVRAAGEYPLDYG 238
 DB 179 SGOQFNRNDQK--GDVAGAHYHLLRAGTNHOGAGNGSYSGYQVRAAGEYPLDYG 237
 QY 239 SKDGSQSPMFIYDAEKQKMLINGILNEGNPEGEKNGFOLVRKSYEDRIFFERDLTSLAT 298
 DB 238 SFGDSQSPMFIYDAEKQKMLINGILNEGNPEGEKNGFOLVRKSYEDRIFFERDLTSLAT 297
 QY 239 RAGNGYTTSGDNG----QGSITQKSGISEITLITLAMSU--PLAKEDKXVHPRIDQ 352
 DB 238 TNAEYRNTGSDHNRVATIKSTLPKCAIQPRYVGLYDYSOLHARDKNGDSRSPYK 357
 QY 353 NIVSPRLNNGETLYFMDQKOSLIFASOINOGAGLYFEGNFTVSPNSNO--TMQAGAHY 411
 DB 358 NIVSPRLNNGETLYFMDQKOSLIFASOINOGAGLYFEGNFTVSPNSNO--TMQAGAHY 417
 QY 412 SENSITVWKNVGEHDLRLSKIGKTLHVQANKENGSIYVGDKEVILLEGADQGNKQAF 471
 DB 418 GEBSTVEWQVHNPEEDRLSKIGKTLHVQANKENGSIYVGDKEVILLEGADQGNKQAF 477
 QY 472 SEIGLVSQKTVQALDDQPTDKYFGFRGRLDNLNHSITKRIQNTDEAMITMHNH 531
 DB 478 KEVGIVSGRATVQALDDQPTDKYFGFRGRLDNLNHSITKRIQNTDEAMITMHNH 537
 QY 532 TOANVTITNGNSIVLPNGKNTKLDYRKEIAYNGFGEPIKJKNHGRNLITYPEETEDR 591
 DB 538 SGTANITITNGNAT---NSDSKQUTKMDLAFNMFGEEDQKAKTNGRLANVYQVNAEN 593
 QY 592 TLLISGGINLKDITQYTKGLFFSGRPTPHAVNHLNKKWSEMEGIPQGEIWMDDHINT 651
 DB 594 HLLISGGINLNGITQNGGTLVFSGRPTPHAVNHLRLDSNMEGIPQGEIWMDDHINT 653
 QY 652 FKAENPQIKGSAVYSRVSSIEBNTVTSNNANATFGVVPQOQNTCTSDMTGLTCKT 711
 DB 654 FKAENPQIKGSAVYSRVSSIEBNTVTSNNANATFGVVPQOQNTCTSDMTGLTCKT 713
 QY 712 VDLPTVYNSIPKQJNGSINLTNDNATAVVGLAKNGVTLTNSOFTLSNNATQIGN 771
 DB 714 VDLPTVYNSIPKQJNGSINLTNDNATAVVGLAKNGVTLTNSOFTLSNNATQIGN 773
 QY 772 IRLSDNSTATVDNANLNGVHLTDSAQFSLKNSHPSHOIQGDKGTVTYLENAWTMPSDT 831
 DB 774 IRLSDNSTATVDNANLNGVHLTDSAQFSLKNSHPSHOIQGDKGTVTYLENAWTMPSDT 833
 QY 832 TLQNLTLNNTITLNSAYASASNTP--RPSLETPTTAHERFTLTVNGKLSGGGT 889
 DB 834 TLQNLTLNNTITLNSAYASASNTP--RPSLETPTTAHERFTLTVNGKLSGGGT 893
 QY 890 POFSSLEFGYKDKLKLSDAEGDYILSVNTKEPETHLEOLTVESKONOPSLDKMET 949
 DB 894 POFSSLEFGYKDKLKLSDAEGDYILSVNTKEPETHLEOLTVESKONOPSLDKMET 953
 QY 950 LENDHVDAGALRYKLVKNDGEFRLANPIKEQELHNDLVYRAQAFRTLEAKQVEPAAQT 1009
 DB 954 LENDHVDAGALRYKLVKNDGEFRLANPIKEQELHNDLVYRAQAFRTLEAKQVEPAAQT 1013
 QY 1010 GEPKTRBRARAAFFDTLPQOSLNLLEAKQELTAETOKSKATYKXVSKRAV--PSD 1067
 DB 1014 GEPKTRBRARAAFFDTLPQOSLNLLEAKQELTAETOKSKATYKXVSKRAV--PSD 1067
 QY 1014 SKARVSR--RAVSDPLPAOSLNLLEAKQEL--LTTEYJUSKA--KXVSKRAVRS 1067

QY 1068 PLLDQSLEFALAEVIDAPQOSEKDLAQEAEK--QRKQKDLISRYNSALSSELSATN 1126
 DB 1068 TLPDQ---LQNALVELIADQOQVKEKQTOGEFEKQRKQKELISRYNSALSSELSATN 1124
 QY 1127 SMLSVDLELRLVYDQASAVNTINADNRHYSDAFPAVQOQCTNRCQVQKXLANGR 1186
 DB 1125 SMLSVDLELRLVYDQASAVNTINADNRHYSDAFPAVQOQCTNRCQVQKXLANGR 1183
 QY 1187 IGAVFSHRSNTPFDOVKNAHTLTMASGPAQYQWGDLOFGVNRGTGISASMAEBSGRX 1246
 DB 1184 IGAVFSHRSNTPFDOVKNAHTLTMASGPAQYQWGDLOFGVNRGTGISASMAEBSGRX 1243
 QY 1244 IHRKAINGVWASVQFRGLQGLIQVYGVNRYPERENYQSEBVRVTPSLAFNRNAGI 1303
 DB 1304 RVDYFTPTDNISVKNYFENVYVYSNANVQTVNLTVIQPFGRVYQKQVGLKEIILHF 1363
 QY 1367 QISAFISKQSGSGLQKQWGVKLGYN 1394
 DB 1364 QISAFISKQSGSGLQKQWGVKLGYN 1391

RESULT 9
 ABU37787
 ID ABU37787 standard; protein; 1449 AA.
 AC ABU37787;
 XX
 AC ABU37787;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #23314.
 XX
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Neisseria meningitidis.
 XX
 PN M020027183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002MO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) BLITRA PHARM INC.
 XX
 XX Wang L, Zarnudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GD, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PESDB; ACA41657.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 PS Claim 25; SEQ ID NO 65711; 1766pp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 623 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding

the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for the gene product or that has an activity against a biological pathway or that inhibits cellular proliferation; (7) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1449 AA:

Query Match 57 48: Score 4178; DB 6; Length 1449:
Best Local Similarity 56.28; Pred. No. 6,7e-232;
Matches 816; Conservative 214; Mismatches 325; Indels 98; Gaps 17;

10 FLTACISLIGVSOAAGHTFGIDYDYPFAKNGKFTVGAQNIKYVKKOQOLVGTSM 69
27 YLALCISFGLLPQAKAGHTFGLNYDYPFAKNGKFAVGAQNIKYVKKOQOLVGTSM 86
70 KAPMIDFSVSNVGAALVENVYIVSAHVAGTIVDPAEKNPDPQHFYTKYKANN 129
87 KAPMIDFSVSNVGAALVENVYIVSAHVAGTIVDPAEKNPDPQHFYTKYKANN 146
130 KCONLHPYEDDTHNRLKFTTAAPIDMTSMNGSTYSRTYPERPVYISGQFPRND 189
147 KPDMSHPYNGDYHMLKFTTAAPIDMTSMNGSTYSRTYPERPVYISGQFPRND 206
190 QDKGDVAGAVHYLTAGNTNORAGNGSYLAGVRAKAGYPLPIAGSKDGSFMT 249
207 DDKRGLSYSGAMLCGTHMAGKNGVYSLSDGVHNDYDPMPIAAGADSGSFM 266
250 YDAERKQMLNGLIBENPFEGKNGFQIVKSYF-DEIFERDLATSLYTRAGVYTTIS 308
267 YDKNKMLNGLNGLIBENPFEGKNGFQIVKSYF-DEIFERDLATSLYTRAGVYTTIS 326
309 GNDKGGSTTQSGIPSEIKITLANNST---FLAKDKVKNRPY--DGNITYSPRLNNGE 363
327 SNNNGTGTVEETNEKVSNEPKLVYALPDESLNETDK--EPYIAAGVNGYRRRLNNGE 384
364 TLVEPQKQKOSLIFASDINOGAGLYFEGNFTVSNSQTMQAGIVSENVSTVTRVNG 423
385 NLSFIIDYNGKILNININOGAGLYFEGNFTVSNSQTMQAGIVSENVSTVTRVNG 444
424 VEHDRLSKIGKGLTHVQKNGKSSISVDGAVYILEQADQKQKAFSEIIVSNGRTV 483
445 VANDRLSKIGKGLTHVQKNGKSSISVDGAVYILEQADQKQKAFSEIIVSNGRTV 504
484 QLNDKQFOTDKFYFSGRGLDNGHSLTPKRIQNTBEGMIVNNTTQAANVTITGNE 543
505 QLNDKQFOTDKFYFSGRGLDNGHSLTPKRIQNTBEGMIVNNTTQAANVTITGNE 564
544 SILVPGNNINKLDRKEIAYNGWFGETDKNKNRNLNLYKPTTDRITLLSGGTNLKG 603
555 SITPSSKININRLNYSKELAYNGWFGETDKNKNRNLNLYKPTTDRITLLSGGTNLNG 624
604 DITVTKKLFSGRGLPPLAIVNHLNKNKSEKGIPOGLVWHDVINTKRNKPEYKGS 663
625 NITVTKKLFSGRGLPPLAIVNHLNKNKSEKGIPOGLVWHDVINTKRNKPEYKGS 684

664 AVASRNVSISBGMWTVSNNAATGVPVNOQNTICTSDMTGLITTCQKVDLITVYINSI 723
685 AINSNNAKVEGOMHLSNNAQAVGVAPHQSHITICTSDMTGLITTCQKVDLITVYINSI 744
724 PPTQINSIMVDMTANVAVKGLANNTVLRHSQFLLSNNAQVGN----- 771
745 TTDLSGNSVLAHDHMLNLTGLATNGNSANGDRRYTVSNATQNGNSLVGNMAQATEN 804
772 -----IRLSDNSTATVDNNLNGVNLITDSNOFLRNS 804
805 QATLNGTASGNASVNSNNAQMSLITSDMAKANSNALANGVSLADKAVTHFPENS 864
805 HSHQIQDQKQITVLENAWTPSPDITLQNTLNNSTITNSAV-----SASSNNT 856
865 RETQLSGNSMDTALHKOSEWTLPSGTGLNMLDVAITLNSAVRHDAQAQTSVSDT 924
857 PR---RSLTETTPPSAHRNTILTVNGQLSQGTFQTSLSGYSKDKLKLSNARSD 913
925 PRRSRRLSLAVPPPSVSRNTLVNGKLNQGTFRPSRLVGYSDKLKLAESSBST 984
914 YILSVRNTGKEPHTLEQLTVESKDNQPLSDKLTLENDHYAGALRYLVRNDGEFRL 973
985 YTLAVNNTGNEPVSLDQLTVDGKNRPSLSENLFTLQNHVDAGAMRYQLKQOEFRL 1044
974 HNPKEQELHNDLVRAE---QAE-----TLAQVEPFAKTQT-GEPRVSRPAPA 1022
1045 HNPVEQELSDPKGKAEKQAEKQKONQSDALIAAGRDAAKTESVAEP-----ARQ 1097
1023 APPDTLPQUSLNNLNAKQAEITAEYKSKAKTKKRSKRAVSPDPLDQSLALEMALE 1082
1098 AGEVAV---GIMQAEKEX-RVQDKQKSAKQREKETPAPTAPF-----BARRA 1144
1083 VIDAPQSEKORLAQAEAKQ-RKQKOLISRYNSALSELATVNSMLSYODELDLFLVD 1141
1145 RRLDFQ-----PQPOPOPOPRDLISRYANSLSFSPSLTNSVPAVQDELDLFLVE 1197
1142 QASAVMTNIAODPKRYSDAFRAVQOQKTNLRQIGVQALANGRIGAVPSHRSRNTFD 1201
1198 DRNNVMTSGIRDTHYHRSQDFRAVQO-TDLRQIGQNLISGAVGLFSHRNTENTD 1256
1202 BOYKRAHATLTMGSPAYQYQGLQGVNNGVGSISAKMAESQSKIKHAKANTGVNASQ 1261
1257 DQIGNSARLASHAVFVGYGIGRFDIGISTGAGSGLSDGIGGIRRVLYHGIQARR 1316
1262 FRIGQLGIOPGVGNRPFBRENYOSEEVRYKTPSLAFNRVAGIRVDYPTPTDNISYK 1321
1317 AGRGGFLEPYTGATRYFQKQDIYEVNINATPGLAFNRVAGIRVDYPTPTDNISYK 1376
1322 PYFPVNVYDVSNANVQTTNLTVLQOPGRYQKQVGLAELIHFQJISAFISKSQSGQLG 1381
1377 PYLSSTYDASGKVRKRVNTVLAQDFGKTRSMWGVNAEIKGFTLSLHAAAKGPOLE 1436
1382 KQNVGVNGLSYRM 1394
1437 AQHSAGIKLGYRM 1449

RESULT 10
AAV56621
ID AAV56621 standard; protein; 1454 AA.
XX AAV56621;
XX
DT 21-FEB-2000 (first entry)
XX
XX Neisseria meningitidis strain ATCC13090 BASB006 protein sequence.
XX
XX Neisseria meningitidis; BASB006; diagnosis; bacterial; infection;
KM vaccine; antibiotic; upper respiratory tract infection; meningitis;
KM invasive bacterial disease; bacteraemia; screening; antibacterial.
XX
OS Neisseria meningitidis.


```

Db 506 LVSGRTVQVQALNADQNFDPDLYFGFGKGLDLNGLSHSPRLQNTDEGAMVHNODKES 565
Qy 536 NTITNGESIVLPNGNNINKLDYRKXIAYNGPGETDKNGKNGRLNLYKPTEDRTLL 595
Db 566 TTTTGNKDLA-TTGN-NSLDSKKIAYNGPGBKDTTKNGKRLNLYQPAADRTLL 623
Qy 596 SGGTNKGDITQTKGLFFSGRPTPAVYHLNKSWESEGIPOGELWDDHIMKRTPKA 655
Db 624 SGGTNLNGNITQTKGLFFSGRPTPAVYHLNDSQKGIPIGELWDDHIMKRTPKA 683
Qy 656 NFOIKGSAVSRVSVSIEGNTVSNANATFVVPQNTITCSDMTGTITQAKDUL 715
Db 684 NFOIKGSAVSRVSVNAYKGDWHLNHAQVGVAPQSHITCTSDMTGLNCKVEKTTI 743
Qy 716 DTKVINSIEKQIINSGINTLNDNAFANVGLAALNGVTLTNHSGFTLSNNAQTGN 771
Db 744 DDKVIALTKYTDISGVNDLADHMLNLTGLATLNGNSANGDTRIVSHNATQNGLSLV 803
Qy 772 -----ILSDNSTATVYDANLNGVHLTDS 796
Db 804 GNAQATFNQATLNGNTSASGNASFNLSDBAVONGSLTSGNAAKAVSHALNGVSLADK 863
Qy 797 AQSLSKNSHPSHQIOGDKTTVLENAATVMPSDTTQNTLANSITTLNAY----- 849
Db 864 AVHFESSRFTQISGGKDTALHLKDEWTLPSGTELGNTLNLNATITLNSAYHDDAQA 923
Qy 850 -SASNNTR-----RSLETETPTSAHFNTLTNGKLSGGTFOFTSLSEGYSD 922
Db 924 QTSACDAPRRSRSRSLSTPPTSVESFNLTITNGKLNQGTFRFSELPGRSD 993
Qy 903 KTKLSDAAGDYILSVNTGKEBELLEQLTVESKNOPLSDKLKTELENDVDAALRY 962
Db 984 KTKLAASSSGTYTLAVNNTGNEPASLEQLTVGEGKONPLSENFNTLLNEHVDAQAMRY 1043
Qy 963 KLVNDGEERLNPKEDEELNDLVRAE---QAER-----TLBAQVEPRTQY-GE 1011
Db 1044 QLIKQGEFRLNPVEOELSDKLGAEKKAQAEKQNAOSLMLTLAGDAVETESVAB 1103
Qy 1012 PKVSRRAAARAPDTLPDQSLNALAEAOALTFETQSKAKTKVRSKRAVSDPLAD 1371
Db 1104 P-----AKQKGENV---GIMAEBSKK-RVQADKXTALAKRAETRPRTTFP--- 1149
Qy 1072 QSLFALMALEVIDAPQOSEKRLDAEAKOKKQKDIISRYNSALSLSATVNSMLSV 1131
Db 1150 -----PARARARDLPQLOPQOPQORDLISRYNSGLSEFSATLNSVFAV 1195
Qy 1132 QDELRLFYDQASAVNTYIAQKGRISDAPRAYQQKTKYLAQIGVQALLANGRIQAV 1191
Db 1196 QDELDRVFAEERNAVMTSGIRDTKHSQDFRAVRQO-TLHQIQMOKNLSSGVGLIF 1254
Qy 1192 SHSRDNTPEBOYQNHATLTMMSGFAQYQMGDLQGVNVTGTSASKMAEBSRKIRKA 1251
Db 1255 SHNRTENTPDGIGNSALAHGAVFGYIDFPYIGISAGAGSSGSLDGGKIRRV 1314
Qy 1252 INGVNNSYQFLGQIGIOPYFGVNRFFIRBNYQSESVRYKTPSLAFNRNAGIRVDT 1311
Db 1315 LHYGIQARYAAGGCGGLEPHIGATRYFYQADYREYVNLATPGALAFRRYAGIKADYS 1374
Qy 1312 FTPTDNTSVKPYFYVNVYDVSNAVQTTNLTLYLQPPGRKWOKEVGLAAELHQSAP 1371
Db 1375 FKPAOHISITPPLSYLSTYDASGVRTRVNTAVLADDFGKSAEWSMAELKGFLLSH 1434
Qy 1372 ISKQSQSLGKQONVGVKLYRM 1394
Db 1435 AAAAKGQDLAAGHSAGIKLGYRM 1457

```

RESULT 15
 AAB25662
 ID AAB25662 standard; protein: 1457 AA.
 AC AAB25662;
 XX

```

Dt 04-DEC-2000 (first entry)
XX
D2 N. meningitidis amino acid sequence orfl-1.pdp SEQ ID NO:1047.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
XX antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; Memb.
XX
OS Neisseria meningitidis.
XX
PN MO200022430-A2.
XX
PD 20-APR-2000.
XX
PE 08-OCT-1999; 99WO-US023573.
XX
PR 09-OCT-1998; 98US-013794P.
XX
PR 30-APR-1999; 99US-0132068P.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Frizer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WP1; 2000-318079/27.
XX
DR N-PSDB; AAB1302.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be used
PT in the diagnosis and treatment of N. meningitidis infection and other
PT Neisserial infections, for example, N.gonorrhoea.
XX
PS Example 1; Page 113; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic proteins
CC from Neisseria genomic sequences. AAB81453 to AAB82414 represent
CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AAB81260 to AAB81303 and AAB25620 to AAB25663 represent Neisseria DNA
CC sequences and their corresponding proteins; AAB81254 to AAB81259 and
CC AAB81304 to AAB81321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AAB81322 to AAB81452 represent
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against Meningococcus B; against all serotypes; and/or against all
CC pathogenic Neisseriae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious Meningococcus B
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions
XX
SQ Sequence 1457 AA;

```

Query Match 56.14; Score 4082; DB 3; Length 1457;
 Best Local Similarity 55.14; Pred. No. 2.3e-226;
 Matches 806; Conservative 218; Mismatches 329; Indels 110; Gaps 19;

```

Qy 10 ELKACISGIVSQAAGHSTYEGIDVOYRDPFAEKKEFTVGAONIKVYNNKQGLVGTSM 69
Db 27 YLAICLSGFLIPQMAAGHTTFGINTYRSDPAKNGKFAVGAQDLEIVYNNKGELVGSMT 86
Qy 70 KAPVIDFVSYSNAGVAAVLENOYIYSVAVNHYGYTDVDFGAEGANNPDQRFYYKIVKNNY 129
Db 87 KAPVIDFVSYSNAGVAAVLENOYIYSVAVNHYGYTDVDFGAEGANNPDQRFYYKIVKNNY 146

```


/ INFORMATION FOR SEQ ID NO: 2:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 1394 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear
 / MOLECULE TYPE: protein
 / SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-10-645-655-2

Query Match 100.0%; Score 7273; DB 12; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKTYFRINFLTACISLGIVSQAMAGHTYFGIDYQYRRPFAENKGFVGAQNIKYKXKQ 60
 DB 1 MKKTYFRINFLTACISLGIVSQAMAGHTYFGIDYQYRRPFAENKGFVGAQNIKYKXKQ 60
 QY 61 GQIVGSMTKAPMIDSVSRNGVAALVENQYIVSAHVAVGYTDVDFGAGANNPDQHRFT 120
 DB 61 GQIVGSMTKAPMIDSVSRNGVAALVENQYIVSAHVAVGYTDVDFGAGANNPDQHRFT 120
 QY 121 YKIVKNNYKKONLHPEYEDYHNPRLHKFTVEAPIDMTSNNGSTYSDETKYPERVRIG 180
 DB 121 YKIVKNNYKKONLHPEYEDYHNPRLHKFTVEAPIDMTSNNGSTYSDETKYPERVRIG 180
 QY 181 SARQFWRNDQDXGDQVAGAYHLLTAGNTNONGAGNGSYLGGDYRKAAGEYGLPIAGSK 240
 DB 181 SARQFWRNDQDXGDQVAGAYHLLTAGNTNONGAGNGSYLGGDYRKAAGEYGLPIAGSK 240
 QY 241 GDSGSPMFTYAEKQKWLINGLRGNPFPEKENGQVLRKSYFDELFEKJHRTSLYTA 300
 DB 241 GDSGSPMFTYAEKQKWLINGLRGNPFPEKENGQVLRKSYFDELFEKJHRTSLYTA 300
 QY 301 GNGVYTSIGNDNGGOSTTQSGIPSEIKITLANMSLPEAKQKYNPRYDGPNIYSPRLN 360
 DB 301 GNGVYTSIGNDNGGOSTTQSGIPSEIKITLANMSLPEAKQKYNPRYDGPNIYSPRLN 360
 QY 361 NGBTIYFMDQKGSLLFASDINGAGGLYFEGNTFVSPTSNTQOGAGIHYSENSYTVTK 420
 DB 361 NGBTIYFMDQKGSLLFASDINGAGGLYFEGNTFVSPTSNTQOGAGIHYSENSYTVTK 420
 QY 421 VNGVHEDRLSKIGKTLHVAKEKNGKSYVGDQKVLLEQDADQGNQASSELGLVSR 480
 DB 421 VNGVHEDRLSKIGKTLHVAKEKNGKSYVGDQKVLLEQDADQGNQASSELGLVSR 480
 QY 481 GTYQVMDKQPTOKETFEGRGRDLNHSITFKRIQNTDEGAMIVENNTQAAVYIT 540
 DB 481 GTYQVMDKQPTOKETFEGRGRDLNHSITFKRIQNTDEGAMIVENNTQAAVYIT 540
 QY 541 GNSIIVPNGNNINKLDYRKEIAYNGMFEETDKKNGKRLIYPTTEEDTLLSGGN 600
 DB 541 GNSIIVPNGNNINKLDYRKEIAYNGMFEETDKKNGKRLIYPTTEEDTLLSGGN 600
 QY 601 LKGDITQTKGLFEFGPPTPHAYNHLNKKWSMGIPOGRIWDDHDMNIRFKAKNFQIK 660
 DB 601 LKGDITQTKGLFEFGPPTPHAYNHLNKKWSMGIPOGRIWDDHDMNIRFKAKNFQIK 660
 QY 661 GGSAYVSRVSSIEGNMTVSNNANATFGVVPQOQNTICTSDMTGLITCKQKADLTDTKY 720
 DB 661 GGSAYVSRVSSIEGNMTVSNNANATFGVVPQOQNTICTSDMTGLITCKQKADLTDTKY 720
 QY 721 NSIPKTIQNGSINLTNNATAVYGLAKNGVNTLTHSOTLSNNATQIGIRISDNTA 780
 DB 721 NSIPKTIQNGSINLTNNATAVYGLAKNGVNTLTHSOTLSNNATQIGIRISDNTA 780
 QY 781 TVDNANKNGVHLTSAQPSLKNSHSHQIQDQKGTIVYLENNTIMRSDTTLNLTAN 840
 DB 781 TVDNANKNGVHLTSAQPSLKNSHSHQIQDQKGTIVYLENNTIMRSDTTLNLTAN 840
 QY 841 STIIVSAYSASNNTPRRSLETTPTTSAHRENTLYVNGLSQGGTPTTSLFGYK 900
 DB 841 STIIVSAYSASNNTPRRSLETTPTTSAHRENTLYVNGLSQGGTPTTSLFGYK 900

QY 901 SDKIKLSNDAEGDYLSVRNTGKPELTLEQLTVESKQNPISDKIKFTLENDHVDAGAL 960
 DB 901 SDKIKLSNDAEGDYLSVRNTGKPELTLEQLTVESKQNPISDKIKFTLENDHVDAGAL 960
 QY 961 RYKLVKNDGFRHNPRIKQELHNDLVRAEOAERTLEAKQVEPTAKTGTGEPKYSRRA 1020
 DB 961 RYKLVKNDGFRHNPRIKQELHNDLVRAEOAERTLEAKQVEPTAKTGTGEPKYSRRA 1020
 QY 1021 PAAFPDTLPDQSLNALAKQAEITLHETQSKAKTKKYSRAVAFSDPLIDSLFALAA 1080
 DB 1021 PAAFPDTLPDQSLNALAKQAEITLHETQSKAKTKKYSRAVAFSDPLIDSLFALAA 1080
 QY 1081 LEVIDAPQSEKDRLOAEBAEKQKODLISRYNSALSELATVSMMLSVODELRLFFV 1140
 DB 1081 LEVIDAPQSEKDRLOAEBAEKQKODLISRYNSALSELATVSMMLSVODELRLFFV 1140
 QY 1141 DQAGSAVNTIADQKRRYDSAFRAYQOQKTMIRQIGVQKALANGRIQAVFHSRSDNTF 1200
 DB 1141 DQAGSAVNTIADQKRRYDSAFRAYQOQKTMIRQIGVQKALANGRIQAVFHSRSDNTF 1200
 QY 1201 DEOVKNHATLTWMSGPAQYQWGDLPFGVNGTGISASMAEJSRKIRRAINYGNASY 1260
 DB 1201 DEOVKNHATLTWMSGPAQYQWGDLPFGVNGTGISASMAEJSRKIRRAINYGNASY 1260
 QY 1261 QPRIGQIGIOPYGVARYEIERENYQSEEVRYKTSIAFNYNAGIRVDYFTPTDNTISV 1320
 DB 1261 QPRIGQIGIOPYGVARYEIERENYQSEEVRYKTSIAFNYNAGIRVDYFTPTDNTISV 1320
 QY 1321 KEFFVNVYDVSNANVQTTNVLTVLQQPFGRYQWEVGLAEIILHFOISAFISKSGSOL 1380
 DB 1321 KEFFVNVYDVSNANVQTTNVLTVLQQPFGRYQWEVGLAEIILHFOISAFISKSGSOL 1380
 QY 1381 GKQQWYKYGKYN 1394
 DB 1381 GKQQWYKYGKYN 1394

RESULT 3
 US-10-080-505-2
 / Sequence 2, Application US/10080505
 / Publication No. US20030073166A1
 / GENERAL INFORMATION:
 / APPLICANT: St. Gene, Joseph W.
 / TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 / FILER REFERENCE: A-59941-1/RT/DC/DBR
 / CURRENT FILING DATE: 2002-02-22
 / PRIOR APPLICATION NUMBER: US 08/296,731
 / PRIOR FILING DATE: 1994-10-25
 / PRIOR APPLICATION NUMBER: US 09/839,996
 / PRIOR FILING DATE: 2001-04-20
 / NUMBER OF SEQ ID NOS: 58
 / SOFTWARE: Patent version 3.1
 / SEQ ID NO 2
 / LENGTH: 1394
 / TYPE: PRP
 / ORGANISM: Haemophilus influenzae
 US-10-080-505-2

Query Match 100.0%; Score 7273; DB 14; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKTYFRINFLTACISLGIVSQAMAGHTYFGIDYQYRRPFAENKGFVGAQNIKYKXKQ 60
 DB 1 MKKTYFRINFLTACISLGIVSQAMAGHTYFGIDYQYRRPFAENKGFVGAQNIKYKXKQ 60
 QY 61 GQIVGSMTKAPMIDSVSRNGVAALVENQYIVSAHVAVGYTDVDFGAGANNPDQHRFT 120
 DB 61 GQIVGSMTKAPMIDSVSRNGVAALVENQYIVSAHVAVGYTDVDFGAGANNPDQHRFT 120
 QY 121 YKIVKNNYKKONLHPEYEDYHNPRLHKFTVEAPIDMTSNNGSTYSDETKYPERVRIG 180
 DB 121 YKIVKNNYKKONLHPEYEDYHNPRLHKFTVEAPIDMTSNNGSTYSDETKYPERVRIG 180

Db 121 YKIVKANNYKONLHPEEDYHNPRLHKEFVTEAAEIDMTSNMNGSTYSDBRTYKBERVRIG 180
 Qy 121 SGRQFMRNDQKGDVAGAYHYLTAAGNTNKGAGNGSVYIGDVPKAEGBEPLRIGSK 240
 Db 121 SGRQFMRNDQKGDVAGAYHYLTAAGNTNKGAGNGSVYIGDVPKAEGBEPLRIGSK 240
 Qy 241 GDSGSPFHYDAEKOKMLINGILREGNPEEGENGPOLVKRSYFDSIFERDLHTSLYTRA 300
 Db 241 GDSGSPFHYDAEKOKMLINGILREGNPEEGENGPOLVKRSYFDSIFERDLHTSLYTRA 300
 Qy 301 GNGVYTTISGNDGQGSITQKSGIPSEIKITLANNGLPLKEKDXNHNRYDGNITYSRLN 360
 Db 301 GNGVYTTISGNDGQGSITQKSGIPSEIKITLANNGLPLKEKDXNHNRYDGNITYSRLN 360
 Qy 361 NGELTYFMODKQGSILFASDINOGAGGLYFENGNFVSNSNQWQAGIHSENSYTWK 420
 Db 361 NGELTYFMODKQGSILFASDINOGAGGLYFENGNFVSNSNQWQAGIHSENSYTWK 420
 Qy 421 VNGVEHRLSKIGKGTJLVQAKGENKSGISVGDGKYLLEQADPOGNKAFSEIGVSGR 480
 Db 421 VNGVEHRLSKIGKGTJLVQAKGENKSGISVGDGKYLLEQADPOGNKAFSEIGVSGR 480
 Qy 481 GTVOLNDKQFDTCKEYFEGFRGSLDNGHSLTFKRIQNTDEGAMIVNENNTQAAVYIT 540
 Db 481 GTVOLNDKQFDTCKEYFEGFRGSLDNGHSLTFKRIQNTDEGAMIVNENNTQAAVYIT 540
 Qy 541 GNEISIVLPNNNNINKLDYRKELIAYNGFEGTDKNKRNRLLYKFTTDEPRLTLLSGTN 600
 Db 541 GNEISIVLPNNNNINKLDYRKELIAYNGFEGTDKNKRNRLLYKFTTDEPRLTLLSGTN 600
 Qy 601 LKGDITQTKGLTFESGRPTPHAVNHLNKSWESEGIPOGEVMDHDMNKRTKAKENFOIK 660
 Db 601 LKGDITQTKGLTFESGRPTPHAVNHLNKSWESEGIPOGEVMDHDMNKRTKAKENFOIK 660
 Qy 661 GGSAYSRNVSSIEGNMTVSNNANAFGVVPOQNTTCTRSBWTSLTTOQKVDLTDKVI 720
 Db 661 GGSAYSRNVSSIEGNMTVSNNANAFGVVPOQNTTCTRSBWTSLTTOQKVDLTDKVI 720
 Qy 721 NSIKRTQJNGSINTLMDNATAVKGLAMNGNTLTNHSOFTLSNNTQJGNIRLSJNSTA 780
 Db 721 NSIKRTQJNGSINTLMDNATAVKGLAMNGNTLTNHSOFTLSNNTQJGNIRLSJNSTA 780
 Qy 781 TVDNANLNGVNLHTSAPSLKNSHFSHQJQDGKGTVLENNATWMPSDITLQNTLANN 840
 Db 781 TVDNANLNGVNLHTSAPSLKNSHFSHQJQDGKGTVLENNATWMPSDITLQNTLANN 840
 Qy 841 STTLNANAYASNNTPRRRSILETETTPSAHRRNTLTNGKLSGQCFOTTSIPEYK 900
 Db 841 STTLNANAYASNNTPRRRSILETETTPSAHRRNTLTNGKLSGQCFOTTSIPEYK 900
 Qy 901 SDKLTANDAGEYILISVANTGKEPETLEOLTJVESKONCPJSDKLTLENDHVDAGAL 960
 Db 901 SDKLTANDAGEYILISVANTGKEPETLEOLTJVESKONCPJSDKLTLENDHVDAGAL 960
 Qy 961 RYKLVKNDGEFRLNPIKEOJLHNDIVRAOERTLEAKOVEPTAKTJGEEKVSRRAA 1020
 Db 961 RYKLVKNDGEFRLNPIKEOJLHNDIVRAOERTLEAKOVEPTAKTJGEEKVSRRAA 1020
 Qy 1021 RAAPFTLTPQOSILNLEKQJAEITLQKSAKTKKVRASRAVSPPLDQSLPALEAA 1080
 Db 1021 RAAPFTLTPQOSILNLEKQJAEITLQKSAKTKKVRASRAVSPPLDQSLPALEAA 1080
 Qy 1081 LEVIDAPOOSEKRLAOEAEKOKKODLSRYSNSALSEJATVNSMLSVDEJDRLEFV 1140
 Db 1081 LEVIDAPOOSEKRLAOEAEKOKKODLSRYSNSALSEJATVNSMLSVDEJDRLEFV 1140
 Qy 1141 DOAGSAWMTIADQKRYDSAPRAYQOQKTNLRQJGVQKALANGRIGAVFHSRSDNTP 1200
 Db 1141 DOAGSAWMTIADQKRYDSAPRAYQOQKTNLRQJGVQKALANGRIGAVFHSRSDNTP 1200
 Qy 1201 DEOVNHAITLTMNSGAOYQKGDJQFGVNWGTGISASQAEQSKTHKRAINYGNWASY 1260
 Db 1201 DEOVNHAITLTMNSGAOYQKGDJQFGVNWGTGISASQAEQSKTHKRAINYGNWASY 1260

Qy 1261 OFRLQGLGIPYFGVARYFERENYQSEEVKTPSLAFNRNAGIRVDYTPETDNISV 1320
 Db 1261 OFRLQGLGIPYFGVARYFERENYQSEEVKTPSLAFNRNAGIRVDYTPETDNISV 1320
 Qy 1321 KPYFFVNVYDVSNANQTTVNLTVIOQFGRYQKQVGLKAEIILHPOISAFKSGSOL 1380
 Db 1321 KPYFFVNVYDVSNANQTTVNLTVIOQFGRYQKQVGLKAEIILHPOISAFKSGSOL 1380
 Qy 1381 GKCQWGVKLGYSR 1394
 Db 1381 GKCQWGVKLGYSR 1394

RESULT 4

US-10-080-505-7
 ; Sequence 7, Application US/10086505
 ; Publication No. US2003073166A1
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph M.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RET/DCF/DHR
 ; CURRENT APPLICATION NUMBER: US/10/080,505
 ; PRIOR FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 1395
 ; TYPE: PR1
 ; ORGANISM: Haemophilus influenzae
 US-10-080-505-7

Query Match 99.6%; Score 7246.5; DB 14; Length 1395;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1393; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

Qy 1 MKTYVRNLNPLTACISIGIVSOAMAGHTYRGIDYQYRDPFAENKKTGVGAKNKKYNNQ 60
 Db 1 MKTYVRNLNPLTACISIGIVSOAMAGHTYRGIDYQYRDPFAENKKTGVGAKNKKYNNQ 60
 Qy 61 GOLVGSATKAPMIDFVSSENQVAAJYENQYIVSAHVNGYTDVDFGABGNNDDQHRT 120
 Db 61 GOLVGSATKAPMIDFVSSENQVAAJYENQYIVSAHVNGYTDVDFGABGNNDDQHRT 120
 Qy 121 YKIVKANNYKONLHPEEDYHNPRLHKEFVTEAAEIDMTSNMNGSTYSDBRTYKBERVRIG 180
 Db 121 YKIVKANNYKONLHPEEDYHNPRLHKEFVTEAAEIDMTSNMNGSTYSDBRTYKBERVRIG 180
 Qy 181 SGRQFMRNDQKGDVAGAYHYLTAAGNTNKGAGNGSVYIGDVPKAEGBEPLRIGSK 240
 Db 181 SGRQFMRNDQKGDVAGAYHYLTAAGNTNKGAGNGSVYIGDVPKAEGBEPLRIGSK 240
 Qy 241 GDSGSPFHYDAEKOKMLINGILREGNPEEGENGPOLVKRSYFDSIFERDLHTSLYTRA 300
 Db 241 GDSGSPFHYDAEKOKMLINGILREGNPEEGENGPOLVKRSYFDSIFERDLHTSLYTRA 300
 Qy 301 GNGVYTTISGNDGQGSITQKSGIPSEIKITLANNGLPLKEKDXNHNRYDGNITYSRLN 360
 Db 301 GNGVYTTISGNDGQGSITQKSGIPSEIKITLANNGLPLKEKDXNHNRYDGNITYSRLN 360
 Qy 361 NGELTYFMODKQGSILFASDINOGAGGLYFENGNFVSNSNQWQAGIHSENSYTWK 420
 Db 361 NGELTYFMODKQGSILFASDINOGAGGLYFENGNFVSNSNQWQAGIHSENSYTWK 420
 Qy 421 VNGVEHRLSKIGKGTJLVQAKGENKSGISVGDGKYLLEQADPOGNKAFSEIGVSGR 480
 Db 421 VNGVEHRLSKIGKGTJLVQAKGENKSGISVGDGKYLLEQADPOGNKAFSEIGVSGR 480
 Qy 481 GTVOLNDKQFDTCKEYFEGFRGSLDNGHSLTFKRIQNTDEGAMIVNENNTQAAVYIT 540

```

Db      481 GTVQNDKQPDIDKFFPFGRGRLDLSLTFKQNTGEMVYNAHTVQANVIT 540
Qy      541 GNEISIVLPNGNNINKLDYRKELAVNGFGEIDKNGKNGRLNIYKPTEDRELLISGGTN 620
Db      541 GNEISIVLPNGNNINKLDYRKELAVNGFGEIDKNGKNGRLNIYKPTEDRELLISGGTN 620
Qy      601 JGDDITVTKGKLPSPGPTPHAVNHLNKRSGEMESJEGEEVWHDNINTEFRANFOIK 660
Db      601 JGDDITVTKGKLPSPGPTPHAVNHLNKRSGEMESJEGEEVWHDNINTEFRANFOIK 660
Qy      661 GSAVAVSRNVSIEGMMVTSNNANATGCVPNQNTICRSPMTGLTCCXKDLDTTVI 720
Db      661 GSAVAVSRNVSIEGMMVTSNNANATGCVPNQNTICRSPMTGLTCCXKDLDTTVI 720
Qy      721 NSIPKTOGNSINLTNATANYKGLAKNGVLTNHSQFTLSNNATQIGNIRLSDNSTA 780
Db      721 NSIPKTOGNSINLTNATANYKGLAKNGVLTNHSQFTLSNNATQIGNIRLSDNSTA 780
Qy      781 TVDNANLNGNVHLTDSAPSLNKHSHSHQIQSGKTTVLENAITWPSDTTQNTIAN 840
Db      781 TVDNANLNGNVHLTDSAPSLNKHSHSHQIQSGKTTVLENAITWPSDTTQNTIAN 840
Qy      841 STITNSAVSASSNNP--RBSLETTTCTSAHRENTLVNKGSGGTPOFTSSJG 898
Db      841 STITNSAVSASSNNP--RBSLETTTCTSAHRENTLVNKGSGGTPOFTSSJG 898
Qy      899 YKSDKLKLSNDAEGDYLISVNTGKEPETLEQLTVESKQNPISDLKFTLENDHVG 958
Db      901 YKSDKLKLSNDAEGDYLISVNTGKEPETLEQLTVESKQNPISDLKFTLENDHVG 960
Qy      959 AIRYKLVKNGDEERLNPKEOEHLNDVPAQDAEFTLENAQVEPTAKCTGEPKVRSR 1018
Db      961 AIRYKLVKNGDEERLNPKEOEHLNDVPAQDAEFTLENAQVEPTAKCTGEPKVRSR 1020
Qy      1019 ARAAFPTLPDQSLNALBAKQALTAETOKSKAKTKKVSRAVSDPLDQSLFALB 1078
Db      1021 ARAAFPTLPDQSLNALBAKQALTAETOKSKAKTKKVSRAVSDPLDQSLFALB 1080
Qy      1079 AALEVITAPQOSEKRLAQEBAKQKQKULISRYSNLSLSTATVNSMLSVODELRL 1138
Db      1081 AALEVITAPQOSEKRLAQEBAKQKQKULISRYSNLSLSTATVNSMLSVODELRL 1140
Qy      1139 FVDQASAVNTNIAQDKERYDSDARAYQOQKNTLQIQVQALANRIGAVSHSRSDN 1198
Db      1141 FVDQASAVNTNIAQDKERYDSDARAYQOQKNTLQIQVQALANRIGAVSHSRSDN 1199
Qy      1199 TDEOVKNHATLTWMSGFAQYOMGDLQGVNNGTGISASKAESQKHRAKINYGVA 1258
Db      1200 TDEOVKNHATLTWMSGFAQYOMGDLQGVNNGTGISASKAESQKHRAKINYGVA 1259
Qy      1259 SYQFRLQGLIQPYFGVNRYPFERENTYQSESVNKTPTSLAFNRNAGIRVDYTFPTNI 1318
Db      1260 SYQFRLQGLIQPYFGVNRYPFERENTYQSESVNKTPTSLAFNRNAGIRVDYTFPTNI 1319
Qy      1315 SYKPYFVNVYDVSNANCTVNTLVLOOPRGYNOKYVGAELTPOISATISKSGS 1378
Db      1320 SYKPYFVNVYDVSNANCTVNTLVLOOPRGYNOKYVGAELTPOISATISKSGS 1379
Qy      1379 QLGKQANVGLDGYRW 1394
Db      1380 QLGKQANVGLDGYRW 1395

```

```

RESULT 5
US-10-080-505-9
; Sequence 3, Application US/10080505
; Publication No. US20030073166A1
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
; CURRENT APPLICATION NUMBER: US/10/080, 505

```

```

; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,396
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1434
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (568)..(568)
; OTHER INFORMATION: The 'Xaa' at location 568 stands for Ser, Gly, Arg, or Cys.
; NAME/KEY: misc_feature
; LOCATION: (1702)..(1702)
; OTHER INFORMATION: "n" at position 1702 can be any base.
US-10-080-505-9

Query Match      84.1%; Score 6115; DB 14; Length 1434;
Best Local Similarity 82.9%; Pred. No. 0;
Matches 1194; Conservative 81; Mismatches 113; Indels 52; Gaps 10;

Qy      1 MKKTVRLNPLFACSGISGIVSOAMAGHTYFGIDYQYRDFRAEKGFYGAQNTKYTKQ 60
Db      1 MKKTVRLNPLFACSGISGIVSOAMAGHTYFGIDYQYRDFRAEKGFYGAQNTKYTKQ 60
Qy      61 GOLVGTSMKAPMIDFVSVRNGVAALVENQYIVSAVANGYTDVDFGAENNDQHRFT 120
Db      61 GOLVGTSMKAPMIDFVSVRNGVAALVENQYIVSAVANGYTDVDFGAENNDQHRFT 120
Qy      61 GOLVGTSMKAPMIDFVSVRNGVAALVENQYIVSAVANGYTDVDFGAENNDQHRFT 120
Db      61 GOLVGTSMKAPMIDFVSVRNGVAALVENQYIVSAVANGYTDVDFGAENNDQHRFT 120
Qy      121 YKIVKNNYTKQNLHPYEDYHNPRLAKFTVTAAPIDKTSNNGSTYSDDRTYPERVIG 180
Db      121 YKIVKNNYTKQNLHPYEDYHNPRLAKFTVTAAPIDKTSNNGSTYSDDRTYPERVIG 180
Qy      181 SGQVFRNDQDQVAVAGAYHLTAGTNRGAGNGSYLAGDVAKAGEYGLPIASGK 240
Db      181 SGQVFRNDQDQVAVAGAYHLTAGTNRGAGNGSYLAGDVAKAGEYGLPIASGK 240
Qy      241 GDGSSPFYTDAREKQKMLINGI.REGNPEGEKENGFOYVAKSYFDIIFERDLHSLYTRA 300
Db      241 GDGSSPFYTDAREKQKMLINGI.REGNPEGEKENGFOYVAKSYFDIIFERDLHSLYTRA 300
Qy      301 GNGYTTISGNDKQGSITQKSGTPESEIKTTLANSPLPKEDVYHNPREDGNTYSPRLN 360
Db      301 GNGYTTISGNDKQGSITQKSGTPESEIKTTLANSPLPKEDVYHNPREDGNTYSPRLN 360
Qy      361 NGETLYFMQKQSLIFASDINQAGALYFEGNFTVSPNSQCTQAGIYVSENSITVTK 420
Db      369 NGKNIHGBDELSITLITTDIDQAGGLYFEDPFLVSEPTKQETRWKAGIHYSEISVTK 418
Qy      421 VNGVEDRLSKIGKTLHVOAKENGKSTISVGDQVILEQACADQGNKQAFSEIGLVGR 480
Db      419 VNGVEDRLSKIGKTLHVOAKENGKSTISVGDQVILEQACADQGNKQAFSEIGLVGR 478
Qy      481 GTVQNDKQPDIDKFFPFGRGRLDLSLTFKQNTGEMVYNAHTVQANVIT 540
Db      479 GTVQNDKQPDIDKFFPFGRGRLDLSLTFKQNTGEMVYNAHTVQANVIT 538
Qy      541 GNEISIVLP--NGNNINKLDYRKELAVNGFGEIDKNGKNGRLNIYKPTEDRELLISGGT 599
Db      539 GNEISIVLP--NGNNINKLDYRKELAVNGFGEIDKNGKNGRLNIYKPTEDRELLISGGT 598
Qy      600 NLKGDITVTKGKLPSPGPTPHAVNHLNKRSGEMESJEGEEVWHDNINTEFRANFOIK 655
Db      599 NLKGDITVTKGKLPSPGPTPHAVNHLNKRSGEMESJEGEEVWHDNINTEFRANFOIK 658
Qy      660 KGSAAVSRNVSIEGMMVTSNNANATGCVPNQNTICRSPMTGLTCCXKDLDTTVI 719
Db      659 KGSAAVSRNVSIEGMMVTSNNANATGCVPNQNTICRSPMTGLTCCXKDLDTTVI 718
Qy      720 NSIPKTOGNSINLTNATANYKGLAKNGVLTNHSQFTLSNNATQIGNIRLSDNSTA 775

```

```

Db 715 IOSTTQINGSTINLNDNAITVINGLAKNGAVTLTNSQFTLSNNATQSGIKLSNEAN 778
Qy 776 ---DNST-----ATVNNANLNGVHLTSAQFS 800
Db 779 AAVNNATLMDVNLADTSNLTNSQATQIGTISLHQQAQAVDANANLNGVHLTSAQFS 838
Qy 801 LKXSHFSHQDQKGTITVLENATVTPSDTTLQNTLNNSTITLNSASSANTP-RR 859
Db 839 LKXSHFSHQDQKGTITVLENATVTPSDTTLQNTLNNSTITLNSASSANTP-RR 898
Qy 860 RSLFETTPSAHRFNLTLYNGKLSGGGTQFTSLFGVSDKLKLSNDAGDYLISVR 919
Db 899 RSLFETTPSAHRFNLTLYNGKLSGGGTQFTSLFGVSDKLKLSNDAGDYLISVR 958
Qy 920 NTGKEPTELEQTLVSKONQPLSDKFTLNDHVDAGALRYLKVNDGEFRLNPIKE 979
Db 959 NTGKEPTELEQTLVSKONQPLSDKFTLNDHVDAGALRYLKVNDGEFRLNPIKE 1018
Qy 980 QELANDLVRAQAKERTLEAKQVEPTAKTQGEPRKVRRAAPAEPTLTPDQSLNALLEA 1039
Db 1019 QELANDLVRAQAKERTLEAKQVEPTAKTQGEPRKVRRAAPAEPTLTPDQSLNALLEA 1075
Qy 1040 KQALTAETQSKAKTKYVRKRAY---FSDPLDQGS-LFALPALEVIDAPOSERKDL 1095
Db 1075 KQALTAETQSKAKTKYVRKRAY---FSDPLDQGS-LFALPALEVIDAPOSERKDL 1135
Qy 1096 AQ-EEAEKQKQKDLISRYNSALSELATVNSKL.SVQDELDR.FVDAQAGAVNTIAD 1154
Db 1136 AQ-EEAEKQKQKDLISRYNSALSELATVNSKL.SVQDELDR.FVDAQAGAVNTIAD 1195
Qy 1155 KRRYSDAFRAYQOQKTNIRQIGVQKALANGRIKAVFSHSDSDTTPBOYKHNATLTKMS 1214
Db 1196 KRRYSDAFRAY-QQKTNIRQIGVQKALANGRIKAVFSHSDSDTTPBOYKHNATLTKMS 1254
Qy 1215 GFAQYQMGDLQFGVNGVGTGISASKAESESRKIHRAIYVGNASVQFRLQGGIOPYFG 1274
Db 1255 GFAQYQMGDLQFGVNGVGTGISASKAESESRKIHRAIYVGNASVQFRLQGGIOPYFG 1314
Qy 1275 VRRYTERENTOSEEVKTPSLAFNRNAGIRVDYTFPTDNI.SVXRYFVNVYDVNSA 1334
Db 1315 VRRYTERENTOSEEVKTPSLAFNRNAGIRVDYTFPTDNI.SVXRYFVNVYDVNSA 1374
Qy 1335 NVOGTWNLTVIQGFPGYQKQVGIKATLHPQISAFISKQSGOLKQONVGVLYGRW 1394
Db 1375 NVOGTWNLTVIQGFPGYQKQVGIKATLHPQISAFISKQSGOLKQONVGVLYGRW 1434

```

RESULT 6

```

US-10-680-505-17
; Sequence 17, Application US/10080505
; Publication No. US20030073166A1
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph M.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RET/DCF/DHR
; CURRENT APPLICATION NUMBER: US/10/080,505
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1411
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-080-505-17

```

Query Match 86.1%; Score 5827.5; DB 14; Length 1411;
 Best Local Similarity 79.5%; Pred. No. 0;
 Matches 1133; Conservative 107; Mismatches 139; Indels 47; Gaps 12;

```

Qy 1 MKKTVRLNPLTACISLGIVSQAMAGHTYFGIDYQYRDFABNKKEFVGAQNTKYNKQ 60
Db 1 MKKTVRLNPLTACISLGIVSQAMAGHTYFGIDYQYRDFABNKKEFVGAQNTKYNKQ 60
Qy 61 GQVGVSTKAPMIDFVSRRNGVAAVENVYIYVAVHGVYTDVDFABKNNDQHRFT 120
Db 61 GQVGVSTKAPMIDFVSRRNGVAAVENVYIYVAVHGVYTDVDFABKNNDQHRFT 120
Qy 121 YKIVKANNKKNNDLHPEYDYNHPLHKEVTEAPLIDNTSNMNGSYSDRKYBERVIG 180
Db 121 YKIVKANNKKNNDLHPEYDYNHPLHKEVTEAPLIDNTSNMNGSYSDRKYBERVIG 180
Qy 181 SGQVPRNDQKQVQVAGAVHYTLAANTHNRGAGNYSTLGGDYPRKAGEPILNGSK 240
Db 181 SGQVPRNDQKQVQVAGAVHYTLAANTHNRGAGNYSTLGGDYPRKAGEPILNGSK 240
Qy 241 GDSGSPMFLYAEKQKXILNGILRSGNPEBKENGFPOLYKSPYEFERD.LHSLYTRA 300
Db 241 GDSGSPMFLYAEKQKXILNGILRSGNPEBKENGFPOLYKSPYEFERD.LHSLYTRA 300
Qy 301 GNGVYITSGNDGQSSITQKSGIPSEKITLANNSLPKECKKVHNPFRYDGNITYSRLN 360
Db 301 GNGVYITSGNDGQSSITQKSGIPSEKITLANNSLPKECKKVHNPFRYDGNITYSRLN 360
Qy 361 NGETIYFNDQKQSLIFASDINGAGCIYFEGNFTVSPNSQWQAGIHSESTVTK 420
Db 361 NGETIYFNDQKQSLIFASDINGAGCIYFEGNFTVSPNSQWQAGIHSESTVTK 420
Qy 421 VNGVEHRLSKIGKSTLHVOAKENKGSISVQDKVTLLEQADQDNQKQASEIGLVGR 480
Db 421 VNGVEHRLSKIGKSTLHVOAKENKGSISVQDKVTLLEQADQDNQKQASEIGLVGR 480
Qy 481 GTVQVNDQKQPTDKFYFGRGRGLDNGHSITFRKQNTDTSBGMTVNTTQAAVNTIT 540
Db 481 GTVQVNDQKQPTDKFYFGRGRGLDNGHSITFRKQNTDTSBGMTVNTTQAAVNTIT 540
Qy 541 GNSSTYLPRGNINKLTVYRKELIYVNGVPEGTDKNNKNGRLNLYKPTEDPILLSGTN 600
Db 541 GNSSTYLPRGNINKLTVYRKELIYVNGVPEGTDKNNKNGRLNLYKPTEDPILLSGTN 600
Qy 601 LKQDITQYKGLFSGRPTPHAVNHLNKSMESEGIPOGHI.VMDHDMNR.FKAKENFOIK 660
Db 601 LKQDITQYKGLFSGRPTPHAVNHLNKSMESEGIPOGHI.VMDHDMNR.FKAKENFOIK 660
Qy 661 GSAVSVSRVSSISGNVTWSSNNANATFVYPRQONTICTSDMTGLITCQKXIDITDKVI 720
Db 661 GSAVSVSRVSSISGNVTWSSNNANATFVYPRQONTICTSDMTGLITCQKXIDITDKVI 720
Qy 721 NSIPKQINGSTINLNDNAITVINGLAKNGAVTLTNSQFTLSNNATQSGIKLSNEAN 780
Db 721 NSIPKQINGSTINLNDNAITVINGLAKNGAVTLTNSQFTLSNNATQSGIKLSNEAN 780
Qy 781 TVDNANLNGVHLTSAQFS.LKXSHFSHQDQKGTITVLENATVTPSDTTLQNTLNN 840
Db 781 TVDNANLNGVHLTSAQFS.LKXSHFSHQDQKGTITVLENATVTPSDTTLQNTLNN 840
Qy 841 STTANSAVSASSANTP-RRSLETTTPSAHRFNLTLYNGKLSGGGTQFTSLFGV 899
Db 841 STTANSAVSASSANTP-RRSLETTTPSAHRFNLTLYNGKLSGGGTQFTSLFGV 899
Qy 899 KSDKLKLSNDAGDYLISVNTGKEPTELEQTLVSKONQPLSDKFTLNDHVDAGA 959
Db 899 KSDKLKLSNDAGDYLISVNTGKEPTELEQTLVSKONQPLSDKFTLNDHVDAGA 959
Qy 959 LRYLVNDGEFRLNPIKEQELANDLVRAQAKERTLEAKQVEPTAKTQGEPRKVRRA 1019
Db 959 LRYLVNDGEFRLNPIKEQELANDLVRAQAKERTLEAKQVEPTAKTQGEPRKVRRA 1019
Qy 1019 ARAAFPTLTPDQSLNALLEAKQALTEPRQSKA-----KTKVSKR---AV 1064
Db 1019 ARAAFPTLTPDQSLNALLEAKQALTEPRQSKA-----KTKVSKR---AV 1064

```


QY 1065 FSDPLDQSL-----FA-----LEALEVIDADPOSEKORLAEAKKORXOD 1108
 DB 1070 LKALVEKETQIDDFANAKLAOLNTQOLEKALAV---ASQEKERBAQOAKRQKQKD 1126
 QY 1109 LISRYNSALSELATVNSMLSVODELRLFVDOASAVMTNADKRRYDPAEFAYOQ 1168
 DB 1127 LISRSNSALSELATVNSMLSVODELRLFVDOASAVMTNADKRRYDPAEFAYOQ 1185
 QY 1169 QKTNRLOIGVQKALANGRIGAVSHSRSDNTFDEOVNKAETLTMSSGPAOYQMGDIQEGV 1228
 DB 1186 QKTNRLOIGVQKALANGRIGAVSHSRSDNTFDEOVNKAETLTMSSGPAOYQMGDIQEGV 1245
 QY 1223 NYGTGISASKMAEESQSKIRKAIYGVNAYSVQFRLGOLGCPYRGNRPFIERENYQSE 1288
 DB 1246 NYGTGISASKMAEESQSKIRKAIYGVNAYSVQFRLGOLGCPYRGNRPFIERENYQSE 1305
 QY 1289 EYVKTPESLAFNRNAGIRVDYFTPTDNTISVXPYFFVNVYDVSNANVQTTVALVLOOP 1348
 DB 1306 EYVKTPESLAFNRNAGIRVDYFTPTDNTISVXPYFFVNVYDVSNANVQTTVALVLOOP 1365
 QY 1349 EGRNQKEVGLKAEILHFOISAFISKSQSQGLKQONVGVKLGTRM 1394
 DB 1366 EGRNQKEVGLKAEILHFOISAFISKSQSQGLKQONVGVKLGTRM 1411

RESULT 7

US-10-080-505-13
 ; Sequence 13, Application US/10083505
 ; Publication No. US20030073166A1
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph M.
 ; TITLE OR INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RET/DCR/DRR
 ; CURRENT APPLICATION NUMBER: US/10/080,505
 ; PRIOR FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-12-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 13
 ; LENGTH: 1436
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 US-10-080-505-13

Query Match 79.8%; Score 5807; DS 14; Length 1436;
 Best Local Similarity 79.3%; Pred. No. 0;
 Matches 1145; Conservative 78; Mismatches 154; Indels 68; Gaps 14;

QY 1 MKKTVFLNLTACISLGIIVSQAMAGHTYFGIDYIYRDFANNGKFTTGAQNIKIVNKQ 60
 DB 1 MKKTVFLNLTACISLGIIVSQAMAGHTYFGIDYIYRDFANNGKFTTGAQNIKIVNKQ 60
 QY 61 GCLVGTSTKAPMIDSVSVSRNGVAVLVEKQYIVSAHVAVGVTVDGAEAGNPPQHRFT 120
 DB 61 GCLVGTSTKAPMIDSVSVSRNGVAVLVEKQYIVSAHVAVGVTVDGAEAGNPPQHRFT 120
 QY 121 YIVGRNNYKKONLHPEYEDYHNPFLKFTVTEAPITDMSNMNGSTYSDETKYPERVIG 180
 DB 121 YIVGRNNYKKONLHPEYEDYHNPFLKFTVTEAPITDMSNMNGSTYSDETKYPERVIG 180
 QY 181 SGROWRNDQD-----KDCYVAGAYHTIANGTNHQRGANGSYISLGDYVRAAGYGLP 235
 DB 181 SGROWRNDQD-----KDCYVAGAYHTIANGTNHQRGANGSYISLGDYVRAAGYGLP 240
 QY 236 LAGSGDGSBPFIYAEKQKYLINGILRSGNPFEGKEKGRQVLRKSYF--DEIFPSDLHT 294
 DB 241 TGVSTGDSGSMFLIDIKQKYLINGVLTGNPFSGAGNGRQVLRKSYFIDNFVZDLP 300
 QY 295 SLYTBAGNGVYITSGNDNGQSGITQ---KSGISEIKITLNMNSLPLKEDKVNRYR-- 345
 DB 295 SLYTBAGNGVYITSGNDNGQSGITQ---KSGISEIKITLNMNSLPLKEDKVNRYR-- 345

DB 301 TELPRSNHNGHSFTSNNGNGTGTQTINKEVSNPQKRTVQLFNEALKKEKX--EPVYAA 358
 QY 350 DGNPTISPRINNGETLYFPQOKQSSLI PASDINGAGAGLYFEGNFTVSNNGTQAGI 409
 DB 359 GGVNAYEPRINNGKNTYFCGRGTGTLTIENNINQAGGLYFEGNFTVSNNGTQAGI 418
 QY 410 HVSNSTTVMKYNVEHRLSKTGKTLHVOAKGKNGGSSIVSDGVVILEQADQSKQ 465
 DB 419 HVGSDSTVMKYNVEHRLSKTGKTLHVOAKGKNGGSSIVSDGVVILEQADQSKQ 478
 QY 470 AESSIGLVSGRGTVOLNDXOPTDKFYFGFRGSLDINGSLTFPR--QNTDEGAMVNH 529
 DB 479 AFKEVGLVSGRATVOLNADQVDPKNTITFGFRGSLDINGSLTFPR--QNTDEGAMVNH 538
 QY 530 NTOGANVITIGNESIYLP--KNNINKLDYRREIAVNGFEGTDKRNKRNGLNLT--YRPT 568
 DB 539 NTOGANVITIGNESITAPSKNNINKLDYRREIAVNGFEGTDKRNKRNGLNLT--YRPT 598
 QY 589 EDRITLLSGTNLKXDITQIKGLFESGRPTPHANHLNKRMEGELDQGEIYVWDHMI 648
 DB 599 EDRITLLSGTNLKXDITQIKGLFESGRPTPHANHLNKRMEGELDQGEIYVWDHMI 656
 QY 649 NRTFKANFOIKGSAVSRVSSIEGNTVSNANATFGVYVNOQNTICTESDWTGLTT 708
 DB 657 TRTFKAMFOIKGSAVSRVSSIEGNTVSNANATFGVYVNOQNTICTESDWTGLTT 726
 QY 709 CQKVDLTDKTVINS--PKQVINGSLNTDVAITANKGLAKINGVTLTNHSOTLSSNATQ 768
 DB 717 CKVVDLTDKTVINS--PKQVINGSLNTDVAITANKGLAKINGVTLTNHSOTLSSNATQ 776
 QY 769 IGRIRLSDNSTATVNNALNGVNHVLDASQPSLKNHSHQIOGDGTVTLENAVTMP 828
 DB 777 TGNQJNSNNAATVNNALNGVNHVLDASQPSLKNHSHQIOGDGTVTLENAVTMP 836
 QY 829 SDITLQNTLNNSTITLNSAVSASSNNTPR--RSLTETPTPTSAERFNTLVNKGSLGQ 887
 DB 837 SDATLQNTLNNSTITLNSAVSASSNNTPR--RSLTETPTPTSAERFNTLVNKGSLGQ 896
 QY 888 GTFOFTSSLPFGYKSNKTLSDNABGYIILSVNTEGKBPTELBOULIYSSKNQPISDKX 947
 DB 897 GTFOFTSSLPFGYKSNKTLSDNABGYIILSVNTEGKBPTELBOULIYSSKNQPISDKX 956
 QY 948 FTLENDVHAGALRYKLVNDGSEFRLLNPYKQELHNDIVRAQAEARTLEAKQVPTAKT 1007
 DB 957 FTLENDVHAGALRYKLVNDGSEFRLLNPYKQELHNDIVRAQAEARTLEAKQVPTAKT 1016
 QY 1008 QTEPKVRSRARAAPEDTLPPQSLNALAEKQKELTAEFTQSKAKTYQVNSKRAVSD 1067
 DB 1017 QTSNARYRSK--RAVFEDTLPPQSLNALAEKQKELTAEFTQSKAKTYQVNSKRAVSD 1071
 QY 1068 FILDQS-----LFALEALEVID 1085
 DB 1072 TLPPQSLQDLVLAQEVPEPTAEKQKAKKVRSKRAAFSDTDLRLSLKVLVEKLEVIN 1131
 QY 1086 APQOESERDLAQEAEKQKOKDLISRYNSALSELATVNSMLSVODELRLFVDOAQS 1145
 DB 1132 AQOQVYKE---PQDEKQKQKOLISRYNSALSELATVNSMLSVODELRLFVDOAQS 1188
 QY 1146 AAVTNIAQDKRRYDSDAFRAYOQKTNLSQIGVQKALANGRIGAVSHSRSDNTFDEQVK 1205
 DB 1189 AAVTNIAQDKRRYDSDAFRAYOQKTNLSQIGVQKALANGRIGAVSHSRSDNTFDEQVK 1247
 QY 1206 NHATLTMSSGPAOYQMGDIQEGVNAVGTGISASKMAEESQSKIRKAIYGVNAYSVQFRLG 1265
 DB 1248 NHATLTMSSGPAOYQMGDIQEGVNAVGTGISASKMAEESQSKIRKAIYGVNAYSVQFRLG 1307
 QY 1266 QIGIOPFYGVNRYFIERENYQSEEVVKTPESLAFNRNAGIRVDYFTPTDNTISVXYFF 1325
 DB 1308 QIGIOPFYGVNRYFIERENYQSEEVVKTPESLAFNRNAGIRVDYFTPTDNTISVXYFF 1367
 QY 1366 VNYVDVSNANVQTTVALVLOOPFGRYQKEVGLKAEILHFOISAFISKSQSQGLKQON 1385
 DB 1368 VNYVDVSNANVQTTVALVLOOPFGRYQKEVGLKAEILHFOISAFISKSQSQGLKQON 1427

QY 1386 VGVKLGXRM 1394
 Db 1428 VGVKLGXRM 1436

RESULT 8

US-10-080-505-15
 ; Sequence 15, Application US/10080505
 ; Publication No. US20030073166A1
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RFT/DCG/DHR
 ; CURRENT APPLICATION NUMBER: US/10/080.505
 ; PRIORITY FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 15
 ; LENGTH: 1391
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 US-10-080-505-15

Query Match 72.9%; Score 5303.5; DB 14; Length 1391;

Best Local Similarity 74.8%; Pred. No. 0;
 Matches 1053; Conservative 109; Mismatches 215; Indels 31; Gaps 15;

QY 1 MKKTVEINFLTACSLGIVSOAMAGHTYEGIDYQYRDRERKGFVYGAQNTKYANKQ 60
 Db 1 MKKTVEINFLTACSLGIVSOAMAGHTYEGIDYQYRDRERKGFVYGAQNTKYANKQ 60
 QY 61 GQVLTSMKAPMIDFSVSRNGVALVENQYIVSVANNGYTDVFGAAGNNPDQRF 120
 Db 61 GEMGTMMKGVPMPLSSMVGAGYSTLISEQLISVANNGYQVDFGAGENNPDQRF 120
 QY 121 YKIVKNNYKCNLHPIEDDYHNPRLHKTVTAPIDTSMNKSSTYSDRTKYERYRIG 180
 Db 121 YKIVKNNYKCNLHPIEDDYHNPRLHKTVTAPIDTSMNKSSTYSDRTKYERYRIG 180
 QY 121 YKIVKNNYKCNLHPIEDDYHNPRLHKTVTAPIDTSMNKSSTYSDRTKYERYRIG 180
 Db 121 YKIVKNNYKCNLHPIEDDYHNPRLHKTVTAPIDTSMNKSSTYSDRTKYERYRIG 180
 QY 181 SGQFWEANDQX--SDQVAGAYHLLTANGTNHOGAGNGSYLGGYVRAAGBYPLT 238
 Db 181 SGQFWEANDQX--SDQVAGAYHLLTANGTNHOGAGNGSYLGGYVRAAGBYPLT 238
 QY 179 SGQFWEANDQX--SDQVAGAYHLLTANGTNHOGAGNGSYLGGYVRAAGBYPLT 237
 Db 179 SGQFWEANDQX--SDQVAGAYHLLTANGTNHOGAGNGSYLGGYVRAAGBYPLT 237
 QY 239 SKGDSGPMETVDAEKOKMLNGLREGNPFEGKENGFOLVKSKYFDELFRDLATSLYT 298
 Db 239 SKGDSGPMETVDAEKOKMLNGLREGNPFEGKENGFOLVKSKYFDELFRDLATSLYT 298
 QY 238 SFGDSSPMETVDAEKOKMLNGLREGNPFEGKENGFOLVKSKYFDELFRDLATSLYT 297
 Db 238 SFGDSSPMETVDAEKOKMLNGLREGNPFEGKENGFOLVKSKYFDELFRDLATSLYT 297
 QY 239 RAGNGVYTTISGDNQ---QGSITQKSGTPELKITLANMSI-PIKEKDYNNPRYDGP 352
 Db 239 RAGNGVYTTISGDNQ---QGSITQKSGTPELKITLANMSI-PIKEKDYNNPRYDGP 352
 QY 238 TNAERENIGSDHNRVACTIKSTLPKKAIOPRIVGLYDMSQLHDAEKEDDESPTXGP 357
 Db 238 TNAERENIGSDHNRVACTIKSTLPKKAIOPRIVGLYDMSQLHDAEKEDDESPTXGP 357
 QY 353 NTVSPRNGETLYMDQKOSLTFASDINOAGSLFPEENFVSPNSVQ--TMQAGATYH 411
 Db 353 NTVSPRNGETLYMDQKOSLTFASDINOAGSLFPEENFVSPNSVQ--TMQAGATYH 411
 QY 358 NTVSPRNGETLYMDQKOSLTFASDINOAGSLFPEENFVSPNSVQ--TMQAGATYH 417
 Db 358 NTVSPRNGETLYMDQKOSLTFASDINOAGSLFPEENFVSPNSVQ--TMQAGATYH 417
 QY 412 SENSTVTKWNGVEHDLRLSKIGTLFVQANKENGKSI-SVGDKYLLEQADQGNKQAF 471
 Db 412 SENSTVTKWNGVEHDLRLSKIGTLFVQANKENGKSI-SVGDKYLLEQADQGNKQAF 471
 QY 418 GEBSTVEQWHPNEPDRISKIGTLFVQANKENGKSI-SVGDKYLLEQADQGNKQAF 477
 Db 418 GEBSTVEQWHPNEPDRISKIGTLFVQANKENGKSI-SVGDKYLLEQADQGNKQAF 477
 QY 472 SEIGVSGAGTYQLNDKQEPDTKYFEGRGRLDNGHSLTKRYQNTDEGANTVNAHT 531
 Db 472 SEIGVSGAGTYQLNDKQEPDTKYFEGRGRLDNGHSLTKRYQNTDEGANTVNAHT 531
 QY 478 KEVIVSGATVQVNSADQVDPNNITFGRGRRLDNGHSLTKRYQNTDEGANTVNAHT 537
 Db 478 KEVIVSGATVQVNSADQVDPNNITFGRGRRLDNGHSLTKRYQNTDEGANTVNAHT 537
 QY 532 TQAAVTTIGNESIYLPKANNINKLDYRKELIAYNGMGSETDKKKNHGRNLITYPEETDR 591
 Db 532 TQAAVTTIGNESIYLPKANNINKLDYRKELIAYNGMGSETDKKKNHGRNLITYPEETDR 591
 QY 538 SQPMNTITIGNAT---NSDSKQVLTNKKDLAFKMGFEDQKAKTKNGLNANNVQPPYNAK 593
 Db 538 SQPMNTITIGNAT---NSDSKQVLTNKKDLAFKMGFEDQKAKTKNGLNANNVQPPYNAK 593
 QY 592 TLLSGTINLKGDTQITQKGLPFGSGRPTPHAYNHLNRKSEKGLPQSEIYVMDHDMINT 651
 Db 592 TLLSGTINLKGDTQITQKGLPFGSGRPTPHAYNHLNRKSEKGLPQSEIYVMDHDMINT 651

Db 594 HLLSGTINLKGDTQITQKGLPFGSGRPTPHAYNHLNRKSEKGLPQSEIYVMDHDMINT 653
 QY 652 FRAENFOIKGSAVASRVSSIEGNWTVSNNAATGVVQNOQNTICTRSDMTGLTTCOK 711
 Db 654 FRAENFOIKGSAVASRVSSIEGNWTVSNNAATGVVQNOQNTICTRSDMTGLTTCOK 713
 QY 712 VDLITTKYINSIKPTQINSINLUDNATANKGLATLNGNTVTLNHSQFTLSNNAQTQIN 771
 Db 714 VDLITTKYINSIKPTQINSINLUDNATANKGLATLNGNTVTLNHSQFTLSNNAQTQIN 773
 QY 772 IRLSDNSIATVDNANNGVHLTDSAGPSIKNSPHSQIQGKGTVTLNNAATWTPST 831
 Db 774 IRLSDNSIATVDNANNGVHLTDSAGPSIKNSPHSQIQGKGTVTLNNAATWTPST 833
 QY 832 TQONTLNNSTVTLNSAYSSANNTP--RRSLRSTETPTSSAHERFTLVNKLSGOQT 889
 Db 834 TQONTLNNSTVTLNSAYSSANNTP--RRSLRSTETPTSSAHERFTLVNKLSGOQT 893
 QY 890 FQFTSLFGYSDKTLKSDNADAGDYTLISVNTGKEBTELOGLTIVESKNDQPLSDTKAT 949
 Db 894 FQFTSLFGYSDKTLKSDNADAGDYTLISVNTGKEBTELOGLTIVESKNDQPLSDTKAT 953
 QY 950 LENDVADGALRYKLVANDGEFRLHNP-KGQELHNDLVABOQERTLEAKQVEPTAKTOT 1009
 Db 954 LENDVADGALRYKLVANDGEFRLHNP-KGQELHNDLVABOQERTLEAKQVEPTAKTOT 1013
 QY 1010 GSKVSRRAARAAPFTLPDQSLINALEAKQELTAFQSKAKTKYKSRV--FSD 1067
 Db 1014 GSKVSRRAARAAPFTLPDQSLINALEAKQELTAFQSKAKTKYKSRV--FSD 1067
 QY 1068 PLIDQSLFALBALEVTIDAPQSEKDLADQBEAK-QRQNDLTSRNSALSELATN 1126
 Db 1068 PLIDQSLFALBALEVTIDAPQSEKDLADQBEAK-QRQNDLTSRNSALSELATN 1124
 QY 1127 SMLSYQDELDELFPDQASAVMTNLAQDKRRYSQSDAFRAYQOQKTNRLQIGYKALANR 1186
 Db 1125 SMLSYQDELDELFPDQASAVMTNLAQDKRRYSQSDAFRAYQOQKTNRLQIGYKALANR 1183
 QY 1187 ICAVFSHRSNDTPDQVKNHATLTMSGPAQIQWQDLOFGVNNGTGISAQRAEVSQRK 1246
 Db 1184 ICAVFSHRSNDTPDQVKNHATLTMSGPAQIQWQDLOFGVNNGTGISAQRAEVSQRK 1243
 QY 1247 HRKAINYGVNASYQRLQSLQIQLYGVNRYFLRENNQSEBYRVTPSLANRYNGI 1306
 Db 1244 HRKAINYGVNASYQRLQSLQIQLYGVNRYFLRENNQSEBYRVTPSLANRYNGI 1303
 QY 1307 RVDYTFPTDNIIVKYEYFFVNTVVDVSNANTVTLNVLVLOQFGRYQKXVGKALILHF 1366
 Db 1304 RVDYTFPTDNIIVKYEYFFVNTVVDVSNANTVTLNVLVLOQFGRYQKXVGKALILHF 1363
 QY 1367 QLSAETISQSQSLGQOQNVGVKLGXRM 1394
 Db 1364 QLSAETISQSQSLGQOQNVGVKLGXRM 1391

RESULT 9

US-10-080-505-11
 ; Sequence 11, Application US/10080505
 ; Publication No. US20030073166A1
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RFT/DCG/DHR
 ; CURRENT APPLICATION NUMBER: US/10/080.505
 ; PRIORITY FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 11

QY	1	MKTVEANFELTACISLGIYSQAMAGHYEPIIDQYVRRPEANCKEYQAKIMIKYKQ	60
QY	2	MKTVEANFELTACISLGIYSQAMAGHYEPIIDQYVRRPEANCKEYQAKIMIKYKQ	60
Db	1	MKTVEANFELTACISLGIYSQAMAGHYEPIIDQYVRRPEANCKEYQAKIMIKYKQ	60
QY	61	GOLVGSMTKAMLIDFVSERBNQVAALVENQYIVSAVHNVGTVDVDCGAGNPPQCHSPT	120
QY	61	GEMIGTMNGQVPMDDLSMRGSGYSTLISEQHLSVAHVGVGDVDMGEMEPQCHAFK	120
Db	121	YKIVKSNATKDONLHPYEDDYENPRLEKPYTEAAPIDMTSNNGSYSDRTKYPERVAIG	180
QY	121	YKIVKSNATKSDRQ---YNDYQHPRLEKPYTEAPLEMSYMDGNHKNENQYPLRVAIG	178
QY	181	SGRQPRNDQX--GDQVAAHNYLTANTNQRKNGNGSYSLGSDYRKKAGEBPLPLAG	23
Db	179	SGRQMKDDNNKTTGDLAVZG--SWLGINTEDEQAGNQTLELNRQVQNMKQKPLPAG	23
QY	239	SKDSSGPFVIYDAEKOMLINGILREKNGPEKENGQYOLVAKSYPEBIEPDLTSLYT	298
Db	238	SFGDSSGPFVIYDKXVKMLINGVLEBNQPAANGSYDITPKOYFGGLINDITANMD	297
QY	299	RAGNYVYITSGNDG-----QGSITQKSGLESEIKYTLANMGL-PLKQKYNHPRYDGP	352
Db	298	TNAEYFNIGSDHNGRVAITIKSTLPEKAIQEPERYGLYNLSQJHADKKNQDESSPYGP	357
QY	353	NYSPELNGEPLFPMQKQGLIFASDINGAGGYEPENYTVSPMSQ--TMQAGLHY	411
Db	358	NPMSPLNHHKSGITFGQGTGTLTENN--NQGAGSYFEGNIVVAKQNNITWQAGVSV	417
QY	412	SENSVYMKVNGVDEHDLSEKISGTYLQVAGENKSGISVDQKYLLEQADPOGKQAF	471
Db	418	GESSTYEWQVHNPEGRDLSEKISGTYLVMGKNLSLSVNGELVLDQADDSQKQAF	477
QY	472	SEIGVSGGYQVQNDKQPTDIDKTFPFGRGRLDNGSLTEKKIONTDEGAMVHNT	531
Db	478	KEVGIVSGATQVQINSADQYDPNNITYEFGRGRLDNGSLTEKKIONTDEGAMIVHNA	537
QY	532	TOAANTTNGESVYLVNGNKNKL DYKEILAYMGPRBDKXNENRLLYXPTEDR	591
Db	538	SOCHNITITGNKTI---NSDSKQJLNKKDIAFEMGPGQDQAKNENRNVNQPAEN	593
QY	592	TLLSGTNLKKQDITOTKGLFPGSRPTPHAYNHLKKRSEKEGIPQSEIYMDHMYRT	651
Db	594	HLLSGTNLNGNITITQGTGLVPSGRPTPHAYNHLKRLDSNVEGIPQSEIYMDHMYRT	655
QY	652	FAENFOIKGSAVSNVSSIBGNVYSNNAATFGVNPQNTCTRSMTGLTTCQK	711
Db	654	FXAENFOIKGSAVSNVSSIEGNVYSNNAATFGVNPQNTCTRSMTGLTTCQK	713
QY	712	VDLTPKXVINSIKPQVQNSINLTDMANANKGLAKNENVTLLDSQFTLSNNAITQEN	773
Db	714	VDLTPKXVINSIPTQVQNSINLTDMANAVNHLGAKNENVTLLDSQFTLSNNAITQEN	773
QY	772	IRLSDNSIATVNNANLNGVNLTDSCFSLKNSHFHQIQGQKGTVYLENATWTPSDQ	831
Db	774	IKLSNNAANTVNNANLNGVNLMDSCFSLKNSHFHQIQGSEDTYMLENATWTPSDQ	833
QY	832	TLCNLTILANSITTLNSAVSASNNTP--RRSLETPTPSAHRPNTLTWNGKLSGGST	889
Db	834	TLCNLTILANSITVTLNSAVSAISNNAPERRRSLETPTPSAHRPNTLTWNGKLSGGST	893
QY	890	FQTSSTLEFYSDKIKXSNABEDYILSVRPTGZPEBTELOJLYLSEKNQPSLTKLFT	944
Db	894	FQTSSTLEFYSDKIKXSNABEDYILSVRPTGZPEBTELOJLYLSEKNQKPLSDKCLFT	953

RESULT 10
US-10-282-122A-65711
Sequence 65711, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELIURA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,135
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/2230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308

QY 130 KXNDL-HPYBDYHNRPLKFTPEAPLDMTSMNNSSTYDEPTKPERVIRISGROFWRN 188
 DB 147 KAGTGHBYGSDYHNRPLKFTPEAPLDMTSMNNSSTYDEPTKPERVIRISGROFWRN 206
 QY 189 DQKAD-----QVAGAYHLLTAGNTHORAGAGSYLGGVAKAGEYEPILJAGSKD 242
 DB 207 DEDERNANESSYHLLASNYMLVGTTFPAQNSGGGVNLGSKTKASPPYFPLTQSGRSD 266
 QY 243 SSSPWFYDAEKQKMLINGILREGNPEPEKENGEPOLVRKSYF-SEIFERDLHTSLYTAG 301
 DB 267 SSSPWFYDAEKQKMLINGILREGNPEPEKENGEPOLVRKSYF-SEIFERDLHTSLYTAG 326
 QY 302 NGVYISGNDNGOSITOK---SGPSEIK---ITLANSLEKEDKVNPRYOGPNY 355
 DB 327 NGVYISGNDNGOSITOK---SGPSEIK---ITLANSLEKEDKVNPRYOGPNY 385
 QY 356 SPRLNNGEYLFYMDQKOSLIFASDINGAGGLYFEGNNTYSPNSNTOGAGIYHNSNS 415
 DB 386 SPRLNNGEYLFYMDQKOSLIFASDINGAGGLYFEGNNTYSPNSNTOGAGIYHNSNS 445
 QY 416 TYWKGNGVEHDLRSKIGKGLHVOAKENGKSIYVGDGKYLLEQOADDQKQAFSEIG 475
 DB 446 TYWKGNGVEHDLRSKIGKGLHVOAKENGKSIYVGDGKYLLEQOADDQKQAFSEIG 505
 QY 476 YVSGGTVOJLNDKQPDIDKTFEGRGRGLDLGHSILTKLQNTDGMATVNHNTQAA 535
 DB 506 YVSGGTVOJLNDKQPDIDKTFEGRGRGLDLGHSILTKLQNTDGMATVNHNTQAA 565
 QY 536 NTYINGESIYLPNGNNIKLDIRKELAVNGMGEYKXKNGRNLNLYEPEDPTLL 595
 DB 566 NTYINGESIYLPNGNNIKLDIRKELAVNGMGEYKXKNGRNLNLYEPEDPTLL 623
 QY 596 SGGTNYKADITQTKGLFFSGSPREHAYNHLNKMSEMEGLPOSEIYWDHDMINTFPAE 655
 DB 624 SGGTNYKADITQTKGLFFSGSPREHAYNHLNKMSEMEGLPOSEIYWDHDMINTFPAE 683
 QY 656 NFOIGGSAVVRNYSSTIGNWTYSNNNAKATPEVYNQONTCTRSWYGLTTCOKDILT 715
 DB 684 NFOIGGSAVVRNYSSTIGNWTYSNNNAKATPEVYNQONTCTRSWYGLTTCOKDILT 743
 QY 716 ETKVINSIPKTOINGSINTDNATANYKGLAKNGVLTJHSGOFTLSNNAOIGN---- 771
 DB 744 ETKVINSIPKTOINGSINTDNATANYKGLAKNGVLTJHSGOFTLSNNAOIGN---- 803
 QY 772 -----TSLDNSTATVDNANLNGNVEJLDS 796
 DB 804 GNAQATFNQATLNGNTSAGSNASENLSDHAWONGSLTISGNAKANVSHSLKNGVYSJADK 863
 QY 797 AQSLSKSHSHOIGODKCTVTLNNAWTMBSTTLNOKTLNASTTLNLSAY----- 849
 DB 864 AQSLSKSHSHOIGODKCTVTLNNAWTMBSTTLNOKTLNASTTLNLSAY----- 923
 QY 850 -SASSNNTPR-----RSLTETPTSAHRENTLTVNGKLSGGTFOPTSLSFGYSD 902
 DB 924 CTSATDAPRRSRBRRLSLSTYPTPSVSRKTLTVNGKLSGGTFOPTSLSFGYSD 983
 QY 903 KXLSNDAEDYLLSTRNTGKEPTELTQTLVSCNOPSIDKLTLENDHVDAGALRY 962
 DB 984 KXLSNDAEDYLLSTRNTGKEPTELTQTLVSCNOPSIDKLTLENDHVDAGALRY 1043
 QY 963 KLVKNDGEPFLNPIEDELNDLVRAE-----QKER-----TLEKQVEYPRAKYOT-GE 1011
 DB 1044 QILRKGFEFLNPIEDELNDLVRAE-----QKER-----TLEKQVEYPRAKYOT-GE 1103
 QY 1012 PAVSRRAADAPFDLTPDQSLNALNLEAKQDELTAETQSKAKTKKVSRAVAFSDPLD 1071
 DB 1104 P-----AFQAGENV---GIMQAEKK-RVQADKOTLAJRKREHFRPATYAF--- 1149
 QY 1072 QSLPRLZALFVTDAPQSEKRLAQEBAKQKQKQDLISYSNSALSEIATVNSMLSV 1131
 DB 1150 -----RARRARRDLPOLQPOPQPOPQDLSISYANSGLSEBSATLNSVAV 1195

QY 1112 QDEIDRLFVDOQAQAVNTNLAQDKKRYSDAPRAYOQKTNLQIOVQALANGRIGAVF 1191
 DB 1196 QDEIDRLFVDOQAQAVNTNLAQDKKRYSDAPRAYOQKTNLQIOVQALANGRIGAVF 1254
 QY 1192 SHSRNDTDEQVKNHATLTMWGGPAXOQMGDLQPGVNVGGLSASIKMAEBSRTHKA 1251
 DB 1255 SHSRNDTDEQVKNHATLTMWGGPAXOQMGDLQPGVNVGGLSASIKMAEBSRTHKA 1314
 QY 1252 INGVNASYOFLQGLQIPYGVNRYFERENTQSEBEVKTPSLAFNRVNAAGIRVDY 1311
 DB 1315 INGVNASYOFLQGLQIPYGVNRYFERENTQSEBEVKTPSLAFNRVNAAGIRVDY 1374
 QY 1312 FTEPDNISVYFEPVNVYVSNANQVTVNLVLOPPGRYKQKQVGLKALILHQSALF 1371
 DB 1375 FTEPDNISVYFEPVNVYVSNANQVTVNLVLOPPGRYKQKQVGLKALILHQSALF 1434
 QY 1372 ISKQSGOLQKQONVGVLYGVR 1394
 DB 1435 ISKQSGOLQKQONVGVLYGVR 1457

RESULT 12
 US-10-282-122A-65027
 Sequence 65027, Application US/10282122A
 Publication No. US0040029129A1
 GENERAL INFORMATION:
 APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zyckind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA.034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PAKX.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 65027
 LENGTH: 1468
 TYPE: PRT
 ORGANISM: Neisseria gonorrhoeae
 US-10-282-122A-65027
 Query Match 56.0%; Score 4072.5; DB 12; Length 1468;
 Best Local Similarity 55.7%; Pred. No. 1,1e-277;

Matches 811; Conservative 209; Mismatches 351; Indels 85; Gaps 17;

```

QY 10 FLTKCSLGI VSGAAGATTGIDYQYRDPENKGTGTAQNIKYNNKQGLVGSMT 69
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 27 YLALCLSGFILPQARAGHTYFGINYYQYRDPENKGTGTAQNIKYNNKQGLVGSMT 86
QY 70 KAPMIDSVSRNGVALVENQYVSYVAHVNGYTDVDFGAGNNPOHREFTYKTYKNRY 129
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 87 KAPMIDSVSRNGVALADQYVSYVAHVNGYNNVDGAGSGNPOHREFTYKTYKNRY 146
QY 130 KK-DNHPEYEDYHNPRLKFTLEAPIDMTSNMNGSTYSDEKTYDERVRLISGRQWRN 188
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 147 KAGTNGHPYGDYHNPRLKFTVDAEVEVMTSYMDGKMYADLKKYDRIRIAGAQWRNS 206
QY 189 DQDKGD-----QVAGAYHYTLAGTTHNQRAGNGYSTLGSDYVRAAGEYGLPIAGSKD 242
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 207 DEDEPNRESSTYHIAASVSWLVGNTFPAQNGGGGTVALSGEKKIHSIPYGLPTGSGFSD 266
QY 243 SGSPMFTYDAEKOKMLINGILREGNPEKENGFOLYRKYEF-DEIFEDDLYTSLYTRAG 301
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 267 SGSPMFTYDAEKOKMLINGVLTGNPTIGSNGFQVAKOMFYDELIFAGDHSVFEPRH 326
QY 302 NGVYTLISGNDNGGGSITOKS---GISSEIK---ITLAMS-PLKEDKVENPRYDGNIT 355
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 327 NGKTFEVDNNGAGKIDAKHGYSLPYRLKXTVQLFNVSLSETPAEPEYHAA-GGVNSY 385
QY 356 SPRLNGETLYEMDQKOSLIFASDINQAGSLYBEGRFTYSNSQYTWQAGNHYEENS 415
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 366 RPLNNGENISFLIDGKELITLNS-INGAGGLYEPGNFTVSPNNRTEWQAGNHSIDGS 445
QY 446 TVTWKAVGVEHDLRSKIGKGLIYVQAKGENKKSISVSGDKVILEQQAQDQENKQAFSEIG 475
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 446 TVTWKNNNGVANDRLSKIGKGLIYVQAKGENQSVSVGDKVILIDQADQKQKQAFSEIG 505
QY 446 LVSGRGTVQVLDNDKQFDTDKYFEGRGSLDNGHSILFKRIQNTDEGATVNNNTQAA 535
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 506 LVSGRGTVQVLDNDQFNPDKLYFEGRGSLDNGHSILSPRIQNTDEGATVNNNQKES 565
QY 536 NVTLTNGESIVLPNNNNINIKLDYKELIANGHGGTDRKNGENRRLTYKPTEDRILL 595
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 566 TVTLTNGKDTL-TTNN-NNLDSKELIANGHGGTDRKNGENRRLTYKPTEDRILL 623
QY 596 SGGINLKGDITQYKGLFPGSGPTPEAHVHLKRWSEMEGIPQGEIYMDHMINRTPKAE 655
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 624 SCGTNNLNGNITQDNKLFPSGRPTPEAHVHLKRWSEMEGIPQGEIYMDHMINRTPKAE 683
QY 656 NFQIKGSSAVSRNVSSTEGNWTYSNNANATFGVEPQONTICTRSWTGTLTQCKYDLT 715
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 684 NFHLQGGQAVSRNVAKEGDMHLSHQAQVAGVAPQSHITICTRSWTGTLTQCKYDLT 743
QY 716 DTKYNSIPKQJNINSINLTDNATAVKGLAKLNNV----- 752
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 744 DOKYIASLSKIDIRGNVSLADHMLVLTGLATLNGNSIAGQDTHVTTRANTONGNISLV 803
QY 753 -----TLTNSQFTLSNNANTQIGNIRLSDNSTATVADVANNANGVHLTDS 796
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 804 GNAQATENQATLNGMTSASDNASPLNNAAVGNSSITLSDNAAANSHSALNGVSLADK 863
QY 797 KQSLKASHSHQIQGDKGTVTLLENATVTPSDTTLQNTLNNSTITLNSAY----- 849
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 864 AVFFENSRFTGKISGGKDTALHMKDSEWTLPSGELGNLNDLNNATITLNSAYHDAAGA 923
QY 850 -SASNNKTR--RSLETEETPTLSAHEFNTLTKNGLSGOSTEPFSSLEGKSDKXK 905
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 924 QYGSADADARRSRSLSTPTPSASRFTLTVNGLSGQSTFPFSELSFGRSKGLK 993
QY 906 LSNDAEGDYILSVNRYKEPETLEQLTVESKQNPISDLKCTLENDHVDAGALRYKLV 965
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 984 LAESSEGYTLTAVNNVTGMEPVSL-EGTVVEGKDMTPLENNLNTLQNHVDAAGMRQGL 1043
QY 966 KNGEERLNPKEDELNDLVRAAEKERTLEAQVEPTAKCYTGEPEYRSKRAARAAR 1025
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1044 KQGEERLNPKEDELNDLVRAAEKERTLEAQVEPTAKCYTGEPEYRSKRAARAAR 1103

```

```

QY 1026 DILPDQSLNLAEKQAE-----LTAETQSKAKTKVRSRRAVFSOPLDQSLFALEA 1079
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1104 NA-TERKASVAREPARQAGENAGIMQAEKEKQVQADK-----DTLAKQREAEETRP 1154
QY 1080 ALEVIDAQQSEKQRLAQGEAEKQ-RKQKDLISYNSLSLSLSTATVNSMLSYODELDL 1138
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1155 ATTAPRRARRRRD-LPQCPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 1213
QY 1139 FVDOAQSAVWNTLADKREYRDSAFRAVOQKNTNLQICVQJALANGRIGAVFSRSDN 1198
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1214 FAEEDRRNVAWVSGIRDTKHYRSQDRAFRQ-Q-TDLAQIMQKGLSSRGVGLTSHARTGN 1272
QY 1199 TPDEQYKQNHATLTMGSPACQYQMDLQFQVNVGTGTSASMAEPOSRLKHKALNTGVA 1258
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1273 TFDQDIGNSARIALAGVFCQYIGIFRFDIGISAGAGSSGSSLDGIRKIRRRVLAHYQIA 1332
QY 1259 SVQFRLQGLQIOPYFGRNRYFERENQSEYVYKTPSLAEVRYAGRYDYFTFTDNI 1316
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1333 RIRAGFGSGFGRPHIGATRYFVQKADRYRENNVNTATPGIAERKRYRAGIKADYSFKRQHI 1392
QY 1319 SVKPYFFVNVVDSNANVQTVNLTVLQCPFGHRYQKQEVGLKAEILHFOISAFISKQGS 1378
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1393 SITPLISLYTDAASQKVRITRVNTAVLADPFGRTRSAEWGVAHKKFTLSLHAAAKXP 1452
QY 1379 QLGKQANQVYLCYRW 1394
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1453 QLEAGHSAGIKLGYRW 1468

RESULT 13
US-10-220-481-73
; Sequence 73, Application US/10220481
; Publication No. US20040110670A1
; GENERAL INFORMATION:
; APPLICANT: ARICO, Maria B., et al.
; TITLE OF INVENTION: Heterologous Expression of Neisseria Proteins
; FILE REFERENCE: CHIR-15883/01US
; CURRENT APPLICATION NUMBER: US/10/220,481
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: PCT/IB01/00452
; NUMBER OF SEQ ID NOS: 633
; SOFTWARE: SeqM199, version 1.02
; SEQ ID NO 73
; LENGTH: 1439
; TYPE: PRP
; ORGANISM: Neisseria meningitidis
US-10-220-481-73

Query Match 55.6%; Score 4044.5; DB 16; Length 1439;
Best Local Similarity 54.7%; Pred. No. 9,8e-276;
Matches 806; Conservative 218; Mismatches 336; Indels 113; Gaps 21;

```


QY 292 LHTSYTBAGNGVYTGNDNGGSGITOK---SGIDSEIK---ITLTMNGLPLKEXKXK 345
 Db 299 THSVYTBPRONGKXSPNDNGKSGINAKNHNHSLPRKKTIVTLNVLSTAEPEY 358
 QY 346 NRPYDGNVYSPRLNNGETLYMDKQSGILFASDINOAGGLYFEGNPFVSPNSQW 405
 Db 359 HAA-GGVNSYRPLNNGENISFLDGGKGLITLNSINQAGGLYFQDPLFVSENNETW 417
 QY 406 GAGIVSEKSTVYTKVNGVEHRLSKIGKGLTHVQAKGNKGSISVDGKVIIEQADQ 465
 Db 413 GAGVHISEDSTVYTKVNGVANDRLSKIGKGLTHVQAKGNKGSISVDGKVIIEQADQ 477
 QY 466 GKQAFSEIGLVSGRTVOLNDKPDYKPPGGELIYNGHSLTFRIQNDDEAM 525
 Db 478 GKQAFSEIGLVSGRTVOLNDKPDYKPPGGELIYNGHSLTFRIQNDDEAM 537
 QY 526 IYVHNTJAAVYITGNESTVLPNGNINKLDRKEIANGVGFETDKNHNGLNLAYK 585
 Db 533 IYVHNDKESTVITLGNKDLA-TGNN-NSLDEKELIANGVGFETDKNHNGLNLAYK 595
 QY 586 PTERETLILSGGJNKGJITTKKLPFGSGRTPHAVNHLKNKSMESGIPQGEIWDH 645
 Db 596 PAEDRILLISGGJNNGNITQTNKLPFGSGRTPHAVNHLNDMSQXGIPQGEIWDH 655
 QY 646 DMINTEKABNFOIKGSAVSRVNSISEGNTVSNNAAPGVVPRQNTICTSDMTG 705
 Db 656 DMINTEKABNFOIKGSAVSRVNSISEGNTVSNNAAPGVVPRQNTICTSDMTG 715
 QY 706 JTCQKVDLDTKVINSLPKQINGSLNLTJNATANVKGKLANGVYTLNHSOFTLSN 765
 Db 716 LNCVEKCTIDDKVIALSLKTDIGSVDLADHNLJLTJLNTLNGNSAAGDRTYVSHN 775
 QY 766 AJOIAR-----TILSNSTATVDMAN 786
 Db 776 AJOIAR-----TILSNSTATVDMAN 835
 QY 787 LAGNHLTDSACSLNSHSHQIQDGKTVTLNATWPSOTLQNTLNNSTILN 846
 Db 836 LAGNHLTDSACSLNSHSHQIQDGKTVTLNATWPSOTLQNTLNNSTILN 895
 QY 847 SAY-----SASNNTR-----RSLSTETPTSAHRPFLTVNKGKSGQTFQF 892
 Db 896 SAYRHDAAAGQGSANDAPRRSRSRSLSTPTPTVESHFNLTVLNKGKSGQTFQF 955
 QY 893 TSSLEGYKSDKLSNDAGGYLSVNTKSEPETTEQTLVESKNQPISSDKLTLEN 952
 Db 956 MELFGYRSDKMLAASSSEGYTLAVNNTNENPASLEQLTVVEGDKKPLSENLFTLON 1015
 QY 953 CHVDAGLRKYLVMNGEFLNPKICEQLHNDLVRAE-----TLKQVYE 1002
 Db 1016 CHVDAGLRKYLVMNGEFLNPKICEQLHNDLVRAE-----TLKQVYE 1075
 QY 1003 PRAKTQT-GEKVRSHBARAAFPDTJPDGSLNMLEAKQELTAETOKSKAKTKVRSK 1061
 Db 1076 AVEKTESVAEP-----ARQAGENV-----GIMQAEERK-RVQADMDTALAKQEAETR 1124
 QY 1062 RAVSDDPLDQSLFALAEVITAPQASKORLAEZEAHOKQKOLISRSNALSSEL 1121
 Db 1125 PTTATFP-----RARRARDPOLQOPOPOPORLISRYANGSLSEF 1167
 QY 1122 SATVMSLTVODELDELFDVDAQASAVNTIADOKARYSDAFRAYOQKTNLROIQVOKA 1181
 Db 1166 SATLNSVPAVODELDELFDVDAQASAVNTIADOKARYSDAFRAYOQKTNLROIQVOKA 1226
 QY 1182 LANGRISAVESHRSRDNFTEDQVZNNHATLTMNSGFAOYQWMDLOFGVNGVIGISAKRAE 1241
 Db 1227 LGSQVGVILFESHRTENTFDQIGNSARLHAGVFGYGDIFVYIGLSAGAFSSGSLSD 1286
 QY 1242 EOSKIRKXKXNGVMAVQRLQGLQYQGVNARYIEKENYQSEVRYVTSILANR 1301
 Db 1287 GIGKIRRRVLAHYG-QARIRAGFGGFIETPHGATRYVQKADRYENVNVIATPELAETR 1346

QY 1302 YNAGIRVYTPETPLNDNISVXKYPFNVDVSNANQVTVNLTYLQOPRGRVQKXVGKA 1361
 Db 1347 YNAGIRVYTPETPLNDNISVXKYPFNVDVSNANQVTVNLTYLQOPRGRVQKXVGKA 1406
 QY 1362 EILFQISAFISKSQSQSLQKQKQVNGKLYRW 1394
 Db 1407 EILFQISAFISKSQSQSLQKQKQVNGKLYRW 1439
 RESULT 14
 US-09-839-996-3
 Sequence 3, Application US/09839996
 Publication No. US20030009010A1
 GENERAL INFORMATION:
 APPLICANT: St. Geme III, Joseph W.
 Falkow, Stanley
 TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
 STREET: 4 Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: California
 COUNTRY: United States
 ZIP: 94111-4187
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/839,996
 FILING DATE: 20-Apr-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/296,791
 FILING DATE: 25-Aug-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Treacartin, Richard F.
 REGISTRATION NUMBER: 31,801
 REFERENCE/DOCKET NUMBER: A-59941/REF/RMS
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 781-1889
 TELEFAX: (415) 398-3249
 TELEX: 910 277299
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1541 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 US-09-839-996-3
 Query Match 23.0%; Score 1673.5; DI 10; Length 1541;
 Best Local Similarity 28.5%; Pred. No. 1.8e-108;
 Matches 480; Conservancy 246; Mismatches 512; Indels 445; Gaps 56;
 QY 6 FRLNPLTACISGLVSOAMAGHTYFGIDYORRPAENKKEFTVGAONIKYNNKQQLV 65
 Db 6 FRLNPLTACISGLVSOAMAGHTYFGIDYORRPAENKKEFTVGAONIKYNNKQQLV 65
 QY 66 TSWTXX-FLIDPSVYS-RNGVALVENQYIVSYAH-NVGTVDVDS-AEGN-----NDQ 117
 Db 66 TSWTXX-FLIDPSVYS-RNGVALVENQYIVSYAH-NVGTVDVDS-AEGN-----NDQ 125
 QY 118 RFL-----YKIVRNVY-----KXONLAPYEDVYHNPRLHKFTVTAAPID-MTS 160
 Db 118 RFL-----YKIVRNVY-----KXONLAPYEDVYHNPRLHKFTVTAAPID-MTS 185
 QY 126 RVSSEKRYTSVEKATYPTKLNKGTYYTTEBQOTQKREEDYTPRLDKFTVTAAPID 185
 Db 126 RVSSEKRYTSVEKATYPTKLNKGTYYTTEBQOTQKREEDYTPRLDKFTVTAAPID 203
 QY 161 NNNGSTYSQKTYPERVYRIGSGQFRRNDQKGDQ-----VAGAYHY 203
 Db 186 SDAQYVNDQKKYPAFAVRLSSGSQF-----YKKGDNTSLILNHEVGNLKLVDGDAYTYG 242

QY 204 TAGTHNGAGAGVSYL-----GVPRAGEYGP.P.IAGSGDGSMPYTAERKQK 257
 Db 243 IAGTYKVNENKGLIGNSKEHSDPFGILLSCDPLTNVAVLGDGSDPLFYREKGMK 302
 QY 258 LINGILREAGPEEGE-NGFOLVRSYFDEIFERDILTSL-----YTAGNG-VYTIIS 309
 Db 303 LFDSYDFMAGYNNKSWQEMWYKSOFTQDVLMKMSAGSLGKTDYMSNSGKTSITIS 362
 QY 310 NDNGQSSITQKSGITSEIKIT.LANGLPLKEKOKYHNPYDGENIYSPALNGETLYEMD 369
 Db 363 GE-----KSLNDVLAD-----GKXP-----HKGSTVF-- 386
 QY 370 QKQSLIFASDINGAGGLYFEGNFTYSPNS-NTWQAGIHVSENGTYWKVNGVZHRD 428
 Db 387 EGSGTLTNANNIDQAGGLFEGDYVKGTSQNTYTKGAGVSAEKGTVTKVHNPQYDR 446
 QY 429 LSKIGKTLHVAQENKSGISVADKVLLEQOADDQONKASEIGLWAGRGVQALND 488
 Db 447 LAKIGKTLIVEGTDNKGSLKVGDC-VILKQINSGG-QHAFASGIVSGRSTVLIND 505
 QY 449 KQPDTEKTFYFGRGRLDNLNHSITKRLQNTBQAMTVNHTTQANVTITGESITVP 548
 Db 506 KQVDPNSIYFGRGRLDNLNHSITKRLQNTBQAMTVNHTTQANVTITGESITVP 565
 QY 549 N----- 549
 Db 566 NITTEYNIAPEDENPYAFRRIDGGQLYALNENTYTAARKASTREBELFRNGESNEN 625
 QY 550 -----GNINKLYRKEIANGMFEGETDKKANGRLNLTYPTEDETLISG 597
 Db 626 MLYKGTSDAKANVMNHNINERMNGFYGE-BEKNKNGNLMTFKGSEQGRLLNG 684
 QY 598 GTNKGDIQOTGKGLFBSGRPTPHANHLAKRSGSMGI-----EGGTYWDHNM 647
 Db 665 GTNKGDIQOTGKGLFBSGRPTPHANHLAKRSGSMGI-----EGGTYWDHNM 736
 QY 648 INETFAENFOIKGSAVNS-RVWSIEGNTVYSNANAFGVVPMQNTICTRSDMTGL 706
 Db 737 INNETFAENFOIKGSAVNS-RVWSIEGNTVYSNANAFGVVPMQNTICTRSDMTGL 794
 QY 707 TTQKVDLTDYVINSIPKTIQINGSINTDNATNANVGLAKNGVTLJNHSQPLSNN 766
 Db 795 VTCITDKLSD-KALNSFNPILRGVNLTESANF-VLGRKANLFGTIGRSGSQ----- 845
 QY 767 TQGNISLDSNSTA-VNMANLNGNMTLDSQFSLKSHSFGIOGKGTVTYENATMT 826
 Db 846 -----VELTENS-----HMLTGN-----SDVQDLDLANGH----- 871
 QY 827 MEADTTLQNLINLNSSTILNSAYSASNNTPRRSRLETETPTSAEHEFNLTJYNGLSG 886
 Db 872 -----IHNSA-----DINSNNVT-----KNTLITVN-SLSG 896
 QY 887 QGTRPQTSLSFGYSKDKLSDAEGDYLLSVNTGKEPTELLOUTLVESKNOPLSDXL 946
 Db 897 NGSPFYLLDLSNKQDQKVVYKSGATGNFTLVADKTEBPNH-NELTLEAKSQ-RDHL 953
 QY 947 KETLENDHVDAGALRYKLVKNDEPRLNPEIKEOHLNDLV----- 987
 Db 954 NVSLVGNITVDLGAWKYKLRVNGERYDLNTP-EVEKKNQVVTNITTPNNIOADVSPV 1011
 QY 988 -----RAQOARTLEAKOVEP----- 1004
 Db 1012 SNNEELARVDAVPYPPAPATPSETTEYVANSKQBSKIVENKBDKETTMAQREVAK 1071
 QY 1005 -----AKTQGE-----PKVSRRAA 1020
 Db 1072 AKSNVANTQUNEVAOSGETKETQTEKETATVEKEKAYETEKTOYEPKVTQSVSP 1131
 QY 1021 RAAPETLIPDS-----LINALEAK-CALTLAETQSKAKTKVNSKRAVSPBLD 1071
 Db 1132 KQOBSETVQPOAPAREKENDPTVNTKEPQSQNTNATDQPAKETS-----SNVEQPVTE 1185

QY 1072 QSLFLEALAEVTDAPQ-----SEKRIAOEAEKQKQK----- 1107
 Db 1186 ST--TVNTGNSVVENENTTPATTOPTVNSSESNKKNHRSVSPVHNPVATSSND 1243
 QY 1108 -----CLISRYNSAELSEA-----TNSMLSYQDELRLLV-DOASAVNTMIAD 1154
 Db 1244 KSTVALCOLTSTVNTNVALSDAPAKAQVALNVAKAVASOHTISQLEKNNBEGYNNWVSNTSM 1303
 QY 1155 KERYSDAFRAYOQOQNTNRLQIGVOKALANG-RIGAVFHSRSHNTDEQVKNHATLTVM 1213
 Db 1304 KNYSSSQYRRFSKSTQ-TQLGMDQITLNNVQDGGVFTYRNSNNNDKATSKN-TLNAV 1361
 QY 1214 SGFAQY-----QKQDLOFVNVGTGILASRAEBOQSKILHKAIVGNVASYOFLQJGT 1269
 Db 1362 NYSKYVADNHW---YIGDILGYGFOFKLQTHNNAKFAHTAQFGH.TAKAKAFNLNFGI 1418
 QY 1270 QPYFVNRVFLERENYQSEEVAVKTPSLANRYNAGIRVDTFPTPNK-SVKRYFPVNV 1329
 Db 1419 TYIVGVWSTLSNADALDQARIKVNPISVKTAFPAQVDLSTY-HLGEFVTLISARY- 1476
 QY 1330 DVSNNVQTVNLTVALQOPFGRYQKEVGLKAEILHFOISAFISKQSQSLGKQDQWGVK 1389
 Db 1477 DANQSGKINWGEYDPAVNVENQQYVNAQKJLKYHNVKLSLIGGLTAKQAEKQKAEIK 1536
 QY 1390 LGY 1392
 Db 1537 LSF 1539

RESULT 15
 US-10-645-655-3
 ; Sequence 3, Application US/10645655
 ; Publication No. US20040063908A1
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Geme III, Joseph W.
 ; Falkow, Stanley
 ; TITLE OF INVENTION: Haemophilus Adherence and Penetration
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Flehr, Hombach, Test, Albritton & Herbert
 ; STREET: 4 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States
 ; ZIP: 94111-4187
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/645,655
 ; FILING DATE: 20-Aug-2003
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/296,791
 ; FILING DATE: 25-Aug-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Treccartin, Richard F.
 ; REGISTRATION NUMBER: 31,801
 ; REFERENCE/DOCKET NUMBER: A-59941/REF/RMS
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 781-1989
 ; TELEFAX: (415) 398-3249
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1541 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 US-10-645-655-3

Query March 23 04; Score 1673.5; DB 12; Length 1541;
 Best Local Similarity 28.58; Pred. No. 1.6e-108;
 Matches 480; Conservative 246; Mismatches 512; Indels 445; Gaps 56;

QY 6 FLNELTACISLIGVSOAMAGCTYDIDYDDEPAENKCKTGAONIVYKQOOLVG 65
 DB 6 FLNELTALVAVALTETEAALVDDVYQIFDEPAENKCKTGAONIVYKQOOLVG 65
 QY 66 TSMYKA-PMIDSVYS-RNGVAALVENOYIVSAH-NVGTYDDVDF-ABGN---NPDH 117
 DB 66 TALPNGIEMIDFVVDVDRKATLILNPGYVGVAVSNGVSELHNGULNMMNNGAAAH 125
 QY 118 EFT-----YIVYKANN-----KCONLHPYEDYANPRILKFTTEALPD-MTS 160
 DB 126 KVSSEENRYSVEKKEEPTKLNKGTVTTEBOQOKREDYVPRLDKFEVVAPEASTA 185
 QY 161 NNNGSTYSDRKYPERVIRIGSGRQFMRRDODKEDQ-----VAGAYAL 203
 DB 186 SSDACTYDDOKTIPAPVRLSGSSQF---IYKGNNTSLINNHEVGNNILKIVGDHITYG 242
 QY 204 TANTHNOGAGNGSYLG-----GVYKAGBYGPLPIAGSKDSDGSPMFIYDAEKOM 257
 DB 243 IAGPYKVNHEHNGULGFQNSKEHSDPKGLISQDPLIVAVLGDSSPLFYDAEKOM 302
 QY 256 LINGLIEBGNFEGKE-NGFQVLRKSYDELFEEDJMTSL-----YTRAGN-YVTISG 309
 DB 303 LFLSGSYDFMAGVKKKMOEMNLYKSQFTKVLKRDASGLISGKTDVSWSSNCKSTITIG 362
 QY 310 NNNGGSIYOKSGIPSEIKITLANNSLPLKEXDXTNPRYDGNIVYFPLNNGETILYMD 369
 DB 363 GE-----KSLAVDLAD-----GKXK-----NCKSVYF-- 386
 QY 370 QKQSLIFASDINQAGSLYFEGNTPVSPNS-NQYQAGLHVSSESTVWYKNGVEHR 428
 DB 387 EESGTLTANNIDQAGGLFPEGDEYVGTSDMTTKAGAVSVAEGKTVMKVHPQDR 446
 QY 429 LSKIGKGLIYCAKGNKSGISVSGKAVILLEQADQOKKCAFSEFISLGSKGTQLND 488
 DB 447 LKIGKGLIYBGTGNKSLKVDGTVLLKQDTNGSG-QHAPASVAVISGSGSTIYVLD 505
 QY 489 KQPDIDKYEFGRGSLDLNGHS-TEFKRIQNTBEGAMTYNNTQOANVTITGNSIYLP 548
 DB 506 KQVDNNSIYFGRGSLDLNNGSLTFPHINIDQAKLVNEMNTNNSITITGESLIDP 565
 QY 549 N----- 549
 DB 566 NTITPYNIDAPEDNPFERRIKDQGOYLNLNNTYVYALRXKASTRELPKNGSNNEN 625
 QY 550 -----GNINKLDYRKELIANGHFGTDRKNGRLULYKPTEDRLILSG 597
 DB 626 WLNGKTSDEAKKNVNMHNNERNKNGFYGE-EEGKNGNLMVYTEKGSQONRFLITG 684
 QY 598 GINLKGLITQKGLFFSGRPTPHAYNHLNKRSEMEGI-----POGEIYWDHW 647
 DB 685 GINLNGDLIYKGLTILSGRPTPH-----RDIAGISSTKDPHAPANNEVVEDW 736
 QY 648 IRRFTKANTFOIKGSAAVVS-RNYSIEGNVTVSNANATGPVNPQNTICTRSDMTGL 706
 DB 737 IRRNKATVYAVTGNASITVSGRNVANITSNITASNVAQVHIGY--KTDGIVCVRSDYTG 794
 QY 707 TTQCKVULDTQYKINSIPKQJNGSINLTDNATAVYKGLAENGVTLNHSQPTLSNA 766
 DB 795 VTCTDRLSD-KALNSFNPTKRGVNLTESANF-VLGKANTFEGTIQSRKNSQ----- 845
 QY 767 TQIGRILNSNTATVDMANLGNVHLTDSQPSLXNSHFSHQIQGDKGTVTLNENATWT 826
 DB 846 -----VALTENS-----HMHUCGN---SDVHQDLANGH----- 871
 QY 827 MESPITLQNTLANNSTIYNSAYSSANNTPRRRSLETETPTTSAEHRNLTJVNGKLSG 886
 DB 872 -----IHLNSA--DNSSNVT-----KXVTLTVN-SLISG 896

QY 887 QGTGTTSLFGKSDKLSNDABGDYILSTRANKKEPPELTQTLVESKNOPLSKL 946
 DB 897 NGSEFTYVLDLSKQGDKXVVKSAATQFTLOVADCTGEPNH-NEILTLPASRQO--RDHL 953
 QY 947 KFTLNDHVDAKALRYKLVKNDEPRLHNPITKQSLHNDV----- 987
 DB 954 NVSLVGNVYDLAKMKYKLRNNGRIDLNP--EYKRRQVDTNTITTPNNTIQADVSPV 1011
 QY 988 -----BASQAEKTLKAEQVEPT----- 1004
 DB 1012 SNNEIARVDEAPVPAPATPSETTETVAENSKQESITVSKQEDATETTAQNRVAK 1071
 QY 1005 -----AKTQGE-----EYVSRRAA 1023
 DB 1072 AKSNVXANTQITNEVVAOSSEETKQTTETATATEKEKAKVETKTOEVKVSQVSP 1131
 QY 1021 RAAPOTLPDS-----LNLALAEK-QAEITAEQKSAKTKYKRRVAFSDPLD 1071
 DB 1132 KQOSEIYQOPAPRENDPTVINKEPQSQNTVTDVIEQPAKETS-----SNVEQVTE 1185
 QY 1072 QSLPAALEALVETIDAPQO-----SEKDLAQSEAEKQKQK----- 1107
 DB 1186 ST-TVNTGNSVVENPEVTTPATQPTVNSSSNKKPKRRRBSVRVPHNVZPATSSND 1243
 QY 1108 -----DLISRYNSGALSLEISA-----TVNSMTSVQDELDRLPV-DQAQSAVNTNAD 1154
 DB 1244 RSTVALCDLITNTNVAVILSDARAKQFVALNKGAVSQHLSOLEMNEQVYVWVSWTSM 1303
 QY 1155 KRRYSDAPRAYQOQKNTLQIGVOKALANG-RIGAVSHSRSDNTPEDQVNNHACTLTM 1213
 DB 1304 NKYVSSQIRREBSKSTV-TQLGMDQTISSNVQGLGFTYKRSNNPDAITSKA-TLAQV 1361
 QY 1214 SGFAQY-----QWGDZQFGVAVGTGISASPMAEQSRKIKHRAINYVNASYQFRLQGLGI 1269
 DB 1362 NFYSKYVADNMH---YGIDIGYKFGQSLQTNHNAKPAHHTAQFGLTAGAFLNGNGI 1418
 QY 1270 QPTVGNRFFIKREYQSEEVRYKTSLSAENRYNAGRVDTTTPPDNISVYEPFVANYV 1329
 DB 1419 TPIVGVRSYLSNADPFDLQARIKVPIISVKYAFQVDSLYT-HUGFSVYPIISARY- 1476
 QY 1330 PVSNAVQTTVNLTVLQOPEGRYQKEVGLKAEILHFOISAPISKQSGSLQKQOONGVK 1389
 DB 1477 DAAQSGKLVNVDYPRAYTEVNOQOINAGLKIKTHNYKLSLIGLTKAKQAEKQKTAELK 1536
 QY 1390 LGY 1392
 DB 1537 LSF 1539

Search completed: June 23, 2004, 17:39:00
 Job time : 67 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 23, 2004, 17:30:13 ; Search time 195 Seconds

(without alignments)
6977.532 Million cell updates/sec

Title: US-10-645-655-2

Perfect score: 7273
Sequence: 1 MKTVFPLNLTNCLISLGV.....SQSQSLGQNVGKLGTRM 1394

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```
1: Pending Parents: AA Main: *
2: /cgn2_6/ptodata/2/paa/PCTUS COMB .pep: *
3: /cgn2_6/ptodata/2/paa/US06 COMB .pep: *
4: /cgn2_6/ptodata/2/paa/US07 COMB .pep: *
5: /cgn2_6/ptodata/2/paa/US08 COMB .pep: *
6: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
7: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
8: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
9: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
10: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
11: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
12: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
13: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
14: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
15: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
16: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
17: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
18: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
19: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
20: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
21: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
22: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
23: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
24: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
25: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
26: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
27: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
28: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
29: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
30: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
31: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
32: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
33: /cgn2_6/ptodata/2/paa/US09 COMB .pep: *
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	7273	100.0	1394	1	PCT-US03-05226-2	Sequence 2, Appli
2	7273	100.0	1394	31	US-10-645-655-2	Sequence 2, Appli
3	7273	100.0	1394	31	US-10-687-046-2	Sequence 2, Appli
4	7246.5	99.6	1395	1	PCT-US03-05226-7	Sequence 7, Appli
5	7246.5	99.6	1395	31	US-10-687-046-7	Sequence 7, Appli
6	6115	84.1	1434	1	PCT-US03-05226-9	Sequence 9, Appli
7	6115	84.1	1434	31	US-10-687-046-9	Sequence 9, Appli
8	5827.5	80.1	1411	1	PCT-US03-05226-17	Sequence 17, Appli
9	5827.5	80.1	1411	31	US-10-687-046-17	Sequence 17, Appli
10	5807	79.8	1436	1	PCT-US03-05226-13	Sequence 13, Appli
11	5807	79.8	1436	31	US-10-687-046-13	Sequence 13, Appli
12	5303.5	72.9	1391	1	PCT-US03-05226-15	Sequence 15, Appli
13	5303.5	72.9	1391	31	US-10-687-046-15	Sequence 15, Appli
14	5297.5	72.8	1391	1	PCT-US03-05226-11	Sequence 11, Appli
15	5297.5	72.8	1391	31	US-10-687-046-11	Sequence 11, Appli
16	4178	57.4	1449	22	US-09-791-537-115250	Sequence 115250, A
17	4178	57.4	1449	28	US-10-282-122A-65711	Sequence 65711, A
18	4090.5	56.2	1454	20	US-09-673-896-2	Sequence 2, Appli
19	4090.5	56.2	1454	32	US-10-742-345-2	Sequence 2, Appli
20	4085	56.2	1457	17	US-09-303-518D-650	Sequence 650, App
21	4085	56.2	1457	22	US-09-791-537-10489	Sequence 40489, A
22	4085	56.2	1457	22	US-09-806-866-87	Sequence 87, App
23	4085	56.2	1457	23	US-09-806-866A-1047	Sequence 1047, Ap
24	4085	56.2	1457	26	US-10-018-470A-87	Sequence 87, Appli
25	4085	56.2	1457	28	US-10-220-481-71	Sequence 71, Appli
26	4083	56.1	1457	20	US-09-673-896-4	Sequence 4, Appli
27	4083	56.1	1457	32	US-10-742-345-4	Sequence 4, Appli
28	4072.5	56.0	1457	22	US-10-282-122A-65027	Sequence 65027, A
29	4072.5	56.0	1457	22	US-09-791-537-110808	Sequence 110808, A
30	4067	55.9	1449	17	US-09-303-518D-652	Sequence 652, App
31	4061.5	55.8	1449	17	US-09-303-518D-654	Sequence 654, App
32	4046	55.6	1455	30	US-10-484-703-32	Sequence 32, Appli
33	4044.5	55.6	1459	28	US-10-220-481-73	Sequence 73, Appli
34	3409.5	46.9	1188	30	US-10-484-703-34	Sequence 34, Appli
35	3352.5	46.1	1136	30	US-10-484-703-36	Sequence 36, Appli
36	2941	40.4	956	30	US-10-484-703-33	Sequence 33, Appli
37	2886	39.7	914	30	US-10-484-703-35	Sequence 35, Appli
38	2347	32.3	979	17	US-09-303-518D-648	Sequence 648, App
39	1676	23.0	1431	20	US-09-674-546A-2601	Sequence 2601, Ap
40	1673.5	23.0	1540	1	US-08-578-788-2	Sequence 2, Appli
41	1673.5	23.0	1541	1	PCT-US03-05226-3	Sequence 3, Appli
42	1673.5	23.0	1541	31	US-10-645-655-3	Sequence 3, Appli
43	1673.5	23.0	1541	31	US-10-687-046-3	Sequence 3, Appli
44	1662.5	22.9	1545	1	PCT-US03-05226-4	Sequence 4, Appli
45	1662.5	22.9	1545	31	US-10-645-655-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
PCT-US03-05226-2
; Sequence 2, Application PC/TUS0305226
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RT/DCR/DHR
; CURRENT APPLICATION NUMBER: PCT/US03/05226
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
PCT-US03-05226-2
Query Match 100.0%; Score 7273; DB 1; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
```

Matches 1394;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MKTYFRLNPLTACISLIGIVSOAMAGHTYGDIDYQYRDFAENKGTGVAQNIKYKQ	60	
DB	1	MKTYFRLNPLTACISLIGIVSOAMAGHTYGDIDYQYRDFAENKGTGVAQNIKYKQ	60	
QY	61	GQVGTSTKTAPIWIDPSVSRNGVAALVENQYIVSVANVGYTDVDFGABNNPDQHEFT	120	
DB	61	GQVGTSTKTAPIWIDPSVSRNGVAALVENQYIVSVANVGYTDVDFGABNNPDQHEFT	120	
QY	121	YKTYKRNKYKKNLHYEDDYNHRLKRTTEAAPIDMTSNMNSTSTRTYPRRPIG	180	
DB	121	YKTYKRNKYKKNLHYEDDYNHRLKRTTEAAPIDMTSNMNSTSTRTYPRRPIG	180	
QY	151	SGROPKNDODKQJYAGAXYLLTAGNTHQKAGNGSYLGGDVAAGYELPIASK	240	
DB	151	SGROPKNDODKQJYAGAXYLLTAGNTHQKAGNGSYLGGDVAAGYELPIASK	240	
QY	241	SDSGSPWFIYDAEKQKWL-MGILBGNPFEGKNGFOLVRKSYEDBIFFEDLHTSLYTRA	300	
DB	241	SDSGSPWFIYDAEKQKWL-MGILBGNPFEGKNGFOLVRKSYEDBIFFEDLHTSLYTRA	300	
QY	301	GNGVYTIISGNDNGQSSITQKSGISEIKITLAMSILPKKCKYHNRYDGRNTSPBLN	360	
DB	301	GNGVYTIISGNDNGQSSITQKSGISEIKITLAMSILPKKCKYHNRYDGRNTSPBLN	360	
QY	361	NGEITLYEYDKQKSLIFASDINOAGAGLYEENFTVSPNSQTOGAGIHVSENSTVWK	420	
DB	361	NGEITLYEYDKQKSLIFASDINOAGAGLYEENFTVSPNSQTOGAGIHVSENSTVWK	420	
QY	421	VNGVEHRLSKTGKTLHYQAKGNGKSIIVSGKTYLBOQADQGNKAFSEISLUSGR	480	
DB	421	VNGVEHRLSKTGKTLHYQAKGNGKSIIVSGKTYLBOQADQGNKAFSEISLUSGR	480	
QY	481	GVVQANDKQPDIDKTFPFRGRJDLNGSLTFKSLQNTGEMINHTQANVTIT	540	
DB	481	GVVQANDKQPDIDKTFPFRGRJDLNGSLTFKSLQNTGEMINHTQANVTIT	540	
QY	541	GNSIVLPNGNINIKLDYKELAVNGFETDKNKNGRLNLYKPTBEDRTLLSGGTN	600	
DB	541	GNSIVLPNGNINIKLDYKELAVNGFETDKNKNGRLNLYKPTBEDRTLLSGGTN	600	
QY	601	LKGDITQTKKLFSSSRPTPEHYNHLKNSSEWEGIPQELVYMHQIMRTKAKNEQIK	660	
DB	601	LKGDITQTKKLFSSSRPTPEHYNHLKNSSEWEGIPQELVYMHQIMRTKAKNEQIK	660	
QY	661	GGSAYVSNVSSIEKMTVSNNAATGCVVPOQNTICTRSDWTGLTTCQKDLDTYVI	720	
DB	661	GGSAYVSNVSSIEKMTVSNNAATGCVVPOQNTICTRSDWTGLTTCQKDLDTYVI	720	
QY	721	NSIPKQINSINLTNATATANKGLAKLNGVTLTNHSGFTLSNNAQIGNRLSDNSTA	780	
DB	721	NSIPKQINSINLTNATATANKGLAKLNGVTLTNHSGFTLSNNAQIGNRLSDNSTA	780	
QY	781	YVDANLNGNHLTDSQSLKNSHFSHQIQSKGTYYTLENATYTPSTTLQULITAN	840	
DB	781	YVDANLNGNHLTDSQSLKNSHFSHQIQSKGTYYTLENATYTPSTTLQULITAN	840	
QY	841	STTLNSAYSSASNNTPRRRSLETCTPTSAERFNTLVNGKLSGGGTQFTSLSFGIK	900	
DB	841	STTLNSAYSSASNNTPRRRSLETCTPTSAERFNTLVNGKLSGGGTQFTSLSFGIK	900	
QY	901	SDKLSNDAGDYILSVN-TGKEPETLEQTLVESKDNQPLSDKLFTELDHYDAAL	960	
DB	901	SDKLSNDAGDYILSVN-TGKEPETLEQTLVESKDNQPLSDKLFTELDHYDAAL	960	
QY	961	RKLYVNGEFLNPIKEOHLANDLVABQARTLBAKQVPTAKTQCEPVSBRBA	1020	
DB	961	RKLYVNGEFLNPIKEOHLANDLVABQARTLBAKQVPTAKTQCEPVSBRBA	1020	
QY	1021	RAAPDTLPDQSLNLMEAKQALTAETOKSKATKVKSKAVSFDPDLDQSLFALBA	1080	
DB	1021	RAAPDTLPDQSLNLMEAKQALTAETOKSKATKVKSKAVSFDPDLDQSLFALBA	1080	

QY	1081	LEVITAPQOSERDLAEEAEKORQKOLISRYNSALSELATVNSMLSVQDELRLV	1140	
DB	1081	LEVITAPQOSERDLAEEAEKORQKOLISRYNSALSELATVNSMLSVQDELRLV	1140	
QY	1141	DQOSAVWNTNADOKRYSDAFRAVQOQKTNLFOIGYQKALANRICAIVPSHRSQNTF	1200	
DB	1141	DQOSAVWNTNADOKRYSDAFRAVQOQKTNLFOIGYQKALANRICAIVPSHRSQNTF	1200	
QY	1201	DEQKHAFLTMSSGFQYQWEDLOFVNVGTGISAKAEQSRKIRKAINGVNAST	1260	
DB	1201	DEQKHAFLTMSSGFQYQWEDLOFVNVGTGISAKAEQSRKIRKAINGVNAST	1260	
QY	1261	QFRLGGLGQIPYGVNRYFIERENYQSEBVRVKTPLAFNRNAGIRVDYFTPTDNISV	1320	
DB	1261	QFRLGGLGQIPYGVNRYFIERENYQSEBVRVKTPLAFNRNAGIRVDYFTPTDNISV	1320	
QY	1321	KPFYFNVYDVSNANVQTTNLTIV-QQPFGRYQKQVGAALIHQISAFISQSSQ	1380	
DB	1321	KPFYFNVYDVSNANVQTTNLTIV-QQPFGRYQKQVGAALIHQISAFISQSSQ	1380	
QY	1381	GKQONVGVLAGYRW	1394	
DB	1381	GKQONVGVLAGYRW	1394	

RESULT 2
 US-10-645-655-2
 Sequence 2, Application US/10645655
 GENERAL INFORMATION:
 APPLICANT: St. Geme III, Joseph W.
 TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Flehr, Hobach, Test, Albritton & Herbert
 STREET: 4 Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: California
 COUNTRY: United States
 ZIP: 94111-4187
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/645,655
 FILING DATE: 20-Aug-2003
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/296,791
 FILING DATE: 25-AUG-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Treacatlin, Richard F.
 REGISTRATION NUMBER: 31,801
 REFERENCE/DOCKET NUMBER: A-59941/PFT/BMS
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 781-1989
 TELEFAX: (415) 398-3249
 TELEX: 910 277999
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1394 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-10-645-655-2

Query Match 100.0%; Score 7273; DB 31; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MKCTVRLNPLTACISIGIVSQAMAGHTYFGIDYQYRDPFAENKGFVGAQNIKYNNQ	60
Db	1	MKCTVRLNPLTACISIGIVSQAMAGHTYFGIDYQYRDPFAENKGFVGAQNIKYNNQ	60
Qy	61	GQVGTSMKAPMIDFVSVRNGVAALVENQYISVAHNVGTTVDVFGAENNDQHRFT	120
Db	61	GQVGTSMKAPMIDFVSVRNGVAALVENQYISVAHNVGTTVDVFGAENNDQHRFT	120
Qy	121	YKIVKNNYKKNLHPYEDDYHNPRLHKFVTEAAPIDMTSNNGSTYSDRKTYPERVRIG	180
Db	121	YKIVKNNYKKNLHPYEDDYHNPRLHKFVTEAAPIDMTSNNGSTYSDRKTYPERVRIG	180
Qy	181	SGRQFMRNDQKQDVAGAHYTLTGNNTNORAGNGSYLGADVRKAGEYGLPIAGSK	240
Db	181	SGRQFMRNDQKQDVAGAHYTLTGNNTNORAGNGSYLGADVRKAGEYGLPIAGSK	240
Qy	241	GDGSPMFIYDAEKQKMLINGILRBNPBEKENGFLVVRKSYFDEIFERDLHTSLYTRA	300
Db	241	GDGSPMFIYDAEKQKMLINGILRBNPBEKENGFLVVRKSYFDEIFERDLHTSLYTRA	300
Qy	301	ANGVYITISGNDNGSGSITOKSGIPSEIKITLANWSLPLKKECKYVHNPRIYDGPNIYSRLN	360
Db	301	ANGVYITISGNDNGSGSITOKSGIPSEIKITLANWSLPLKKECKYVHNPRIYDGPNIYSRLN	360
Qy	361	NGEITVPMQKQSLIFASDINOGAGGLYFBEENFTVSNNOGAGIHVSNSSTYTWK	420
Db	361	NGEITVPMQKQSLIFASDINOGAGGLYFBEENFTVSNNOGAGIHVSNSSTYTWK	420
Qy	421	VNGVEHRLSKIGKGLTHVQAKENGNGSIVGDGVLLIEQADQDNKQASEHIGLVSQ	480
Db	421	VNGVEHRLSKIGKGLTHVQAKENGNGSIVGDGVLLIEQADQDNKQASEHIGLVSQ	480
Qy	481	GTVOANDKQFDTFEYFEGGGLDLNGHSLTFKRLONTDEGAMVNTTQOANYTIT	540
Db	481	GTVOANDKQFDTFEYFEGGGLDLNGHSLTFKRLONTDEGAMVNTTQOANYTIT	540
Qy	541	GNESITVPMGNANNKLDYRKEIAYNGMFBETDKNNKMRNLILYKTEJEDRLTLLSGGIV	600
Db	541	GNESITVPMGNANNKLDYRKEIAYNGMFBETDKNNKMRNLILYKTEJEDRLTLLSGGIV	600
Qy	601	LKGIITCTGKCLFESGRPTPHAVNHLNKRSMBSGIPGELVMDHMINIKRIFKLENOIK	660
Db	601	LKGIITCTGKCLFESGRPTPHAVNHLNKRSMBSGIPGELVMDHMINIKRIFKLENOIK	660
Qy	661	GSAVVSBNVSSIEGNVTVSNNAATFVVVPMQOKTICTSDMTGLTTCQKVDLDTKVI	720
Db	661	GSAVVSBNVSSIEGNVTVSNNAATFVVVPMQOKTICTSDMTGLTTCQKVDLDTKVI	720
Qy	721	NSIPKQINGSLNLTDAATAVKGLAKLANSNTLTNHSOFTLSNNAIOGIRLSDNSTA	780
Db	721	NSIPKQINGSLNLTDAATAVKGLAKLANSNTLTNHSOFTLSNNAIOGIRLSDNSTA	780
Qy	781	TYDNANLNGVHLTSAOFLKNSHSEHQOGDKGTVTLLENATMTMPEDTLQWLTIAN	840
Db	781	TYDNANLNGVHLTSAOFLKNSHSEHQOGDKGTVTLLENATMTMPEDTLQWLTIAN	840
Qy	841	STILLNSAVSASSNTPRRSLTETETPTSAHRFNLTJVNGKLSQGGFTGTSLSFGYK	900
Db	841	STILLNSAVSASSNTPRRSLTETETPTSAHRFNLTJVNGKLSQGGFTGTSLSFGYK	900
Qy	901	SDKIKLSNDAEGYILSVRNGKEPTELTDLTVESKXNGLPLSKIKLENDHVDGAL	960
Db	901	SDKIKLSNDAEGYILSVRNGKEPTELTDLTVESKXNGLPLSKIKLENDHVDGAL	960
Qy	961	RYKIVKNDGFRLLNPIKEOELHNDLVRAOAEFTLAKOYEPFAKTGCEPKRSRAA	1020
Db	961	RYKIVKNDGFRLLNPIKEOELHNDLVRAOAEFTLAKOYEPFAKTGCEPKRSRAA	1020
Qy	1021	RAAPBDTLPPQSLNALBAKQALTAETOKSFAKTKYVRSKRAVFSDDLQSLPALAA	1080
Db	1021	RAAPBDTLPPQSLNALBAKQALTAETOKSFAKTKYVRSKRAVFSDDLQSLPALAA	1080

Qy	1081	LEVIDAPQSSKRLAQBFAEKORQKDLISRYSNLSLSELSATVNSMLSVQDELRLFLV	1140
Db	1081	LEVIDAPQSSKRLAQBFAEKORQKDLISRYSNLSLSELSATVNSMLSVQDELRLFLV	1140
Qy	1141	DQASAVNTNIAODKRRYDSAFRAYOQOKTNLRQIGVQKALANGRLGAVFSSRSNDTF	1200
Db	1141	DQASAVNTNIAODKRRYDSAFRAYOQOKTNLRQIGVQKALANGRLGAVFSSRSNDTF	1200
Qy	1201	DEQVNRATLTMSGFAOYQWGDILQFGVNVGTGISASNAEBSQSKLHRAIINYGNASY	1260
Db	1201	DEQVNRATLTMSGFAOYQWGDILQFGVNVGTGISASNAEBSQSKLHRAIINYGNASY	1260
Qy	1261	QRLGQLGIGYRGVNRVFIENYQSEFVRVTPSLANRNAGIRVDTPTPTNISV	1320
Db	1261	QRLGQLGIGYRGVNRVFIENYQSEFVRVTPSLANRNAGIRVDTPTPTNISV	1320
Qy	1321	KPYEFVNTVDVSNANVQITVNLTVLQOPFGYQKQKGLKALILFOISAFISKSQSQOL	1380
Db	1321	KPYEFVNTVDVSNANVQITVNLTVLQOPFGYQKQKGLKALILFOISAFISKSQSQOL	1380
Qy	1381	GKQNVGVKLAGYEM	1394
Db	1381	GKQNVGVKLAGYEM	1394

RESULT 3
US-10-687-046-2
; Sequence 2, Application US/10687046
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RET/DCE/DR
; CURRENT APPLICATION NUMBER: US/10/687,046
; PRIOR FILING DATE: 2003-10-15
; PRIOR FILING DATE: 2002-02-22
; PRIOR FILING DATE: 2002-02-22
; PRIOR FILING DATE: 1994-10-25
; PRIOR FILING DATE: 1994-10-25
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent version 3.1
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-687-046-2

Query Match 100.0%; Score 7273; DB 31; Length 1394;
Beet local similarity 100.0%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MKCTVRLNPLTACISIGIVSQAMAGHTYFGIDYQYRDPFAENKGFVGAQNIKYNNQ	60
Db	1	MKCTVRLNPLTACISIGIVSQAMAGHTYFGIDYQYRDPFAENKGFVGAQNIKYNNQ	60
Qy	61	GQVGTSMKAPMIDFVSVRNGVAALVENQYISVAHNVGTTVDVFGAENNDQHRFT	120
Db	61	GQVGTSMKAPMIDFVSVRNGVAALVENQYISVAHNVGTTVDVFGAENNDQHRFT	120
Qy	121	YKIVKNNYKKNLHPYEDDYHNPRLHKFVTEAAPIDMTSNNGSTYSDRKTYPERVRIG	180
Db	121	YKIVKNNYKKNLHPYEDDYHNPRLHKFVTEAAPIDMTSNNGSTYSDRKTYPERVRIG	180
Qy	181	SGRQFMRNDQKQDVAGAHYTLTGNNTNORAGNGSYLGADVRKAGEYGLPIAGSK	240
Db	181	SGRQFMRNDQKQDVAGAHYTLTGNNTNORAGNGSYLGADVRKAGEYGLPIAGSK	240
Qy	241	GDGSPMFIYDAEKQKMLINGILRBNPBEKENGFLVVRKSYFDEIFERDLHTSLYTRA	300
Db	241	GDGSPMFIYDAEKQKMLINGILRBNPBEKENGFLVVRKSYFDEIFERDLHTSLYTRA	300

QY 301 GNGVTTISGNDGQSLTOKSGIPSEKITLANMSPLKXKDVHNPIDGNITSPRLN 360
 DB 301 GNGVTTISGNDGQSLTOKSGIPSEKITLANMSPLKXKDVHNPIDGNITSPRLN 360
 QY 361 NGELTYPMQKQKSLIFASDINOGAGLYFEENFTVPSNSNOTMOGAGIHVSNSITVTK 420
 DB 361 NGELTYPMQKQKSLIFASDINOGAGLYFEENFTVPSNSNOTMOGAGIHVSNSITVTK 420
 QY 421 VNGVEHDLRSKIGKGLTHVQAKGENKGSISVGDGKVLLEQADDOGNQAFSEIGVSGR 480
 DB 421 VNGVEHDLRSKIGKGLTHVQAKGENKGSISVGDGKVLLEQADDOGNQAFSEIGVSGR 480
 QY 481 GTVQANDKQPTDKFYFGGRGLDNGHSLTFKRIQNTDEGAMIVHNTTOAANVTIT 540
 DB 481 GTVQANDKQPTDKFYFGGRGLDNGHSLTFKRIQNTDEGAMIVHNTTOAANVTIT 540
 QY 541 GNESTIVLPNGNNINKLDYRKELIANGMFGETDKNKNRNLILYKTTEDRTLLSGGN 600
 DB 541 GNESTIVLPNGNNINKLDYRKELIANGMFGETDKNKNRNLILYKTTEDRTLLSGGN 600
 QY 601 LKGDITQTKGLFESGRTPHAVNHLNKSSEMGIPQGETIWDHDMINRTFKAEFQIK 660
 DB 601 LKGDITQTKGLFESGRTPHAVNHLNKSSEMGIPQGETIWDHDMINRTFKAEFQIK 660
 QY 661 GSAVNSRANSSIEGNFTVSNNAATFGVVPNOQKICTSDPTGLTTQKRDLTPTKVI 720
 DB 661 GSAVNSRANSSIEGNFTVSNNAATFGVVPNOQKICTSDPTGLTTQKRDLTPTKVI 720
 QY 721 NSIPKTOJNGSLTJDNATAVKGLANMANTLTJNHSOFTLSNNTOIGNRLSDNSTA 780
 DB 721 NSIPKTOJNGSLTJDNATAVKGLANMANTLTJNHSOFTLSNNTOIGNRLSDNSTA 780
 QY 781 TVDNANINGVHLTJSAQFSLKNSHFSHQIQDCKTIVTLENAITWMTSDITLLQNTLAN 840
 DB 781 TVDNANINGVHLTJSAQFSLKNSHFSHQIQDCKTIVTLENAITWMTSDITLLQNTLAN 840
 QY 841 STITNSAVSASNNTPRRSLTETPTPSAHRPNLTJNGLSQGFQGTSLSGYK 900
 DB 841 STITNSAVSASNNTPRRSLTETPTPSAHRPNLTJNGLSQGFQGTSLSGYK 900
 QY 901 SDKILSNDAGGYILSVRNTGKEPETLEQTLTVESKONCPLEDKITLENDHVDAGL 960
 DB 901 SDKILSNDAGGYILSVRNTGKEPETLEQTLTVESKONCPLEDKITLENDHVDAGL 960
 QY 961 RYKLVKNDGSEFLNPIKEQLENDLYRABOAKERTLEKQVEPTACTOTGEPKYSRRA 1020
 DB 961 RYKLVKNDGSEFLNPIKEQLENDLYRABOAKERTLEKQVEPTACTOTGEPKYSRRA 1020
 QY 1021 RAAPDCTLPPOSILNALAKOAELETAETOKSAKTKYRSRAVFSDDLDSQJFALBA 1080
 DB 1021 RAAPDCTLPPOSILNALAKOAELETAETOKSAKTKYRSRAVFSDDLDSQJFALBA 1080
 QY 1081 LEVIDAPQSEKDRLOAEAEKOKKODLISRYNSALSELASATVSMLSVDELEIRFV 1140
 DB 1081 LEVIDAPQSEKDRLOAEAEKOKKODLISRYNSALSELASATVSMLSVDELEIRFV 1140
 QY 1141 DOAGSAVNTIAODKRRITDSAPRAYOQKTNLROIGYQALANGELIGAVSHSNTPT 1200
 DB 1141 DOAGSAVNTIAODKRRITDSAPRAYOQKTNLROIGYQALANGELIGAVSHSNTPT 1200
 QY 1201 DEOVYKNAITLTMASGFAOYQMSDQFCGVNNGISASISMAEGRKIRKALINQVASY 1260
 DB 1201 DEOVYKNAITLTMASGFAOYQMSDQFCGVNNGISASISMAEGRKIRKALINQVASY 1260
 QY 1261 QFRIGOLGIPYFGVNYFIERENYQSEEVRYKTPSLAFNRKYNAGRVDYTFPTDNISV 1320
 DB 1261 QFRIGOLGIPYFGVNYFIERENYQSEEVRYKTPSLAFNRKYNAGRVDYTFPTDNISV 1320
 QY 1321 KPYEFVNVVSNANOTVNLTVLQOPFERWQKEVGLKELIHQISAFISQSGSL 1380
 DB 1321 KPYEFVNVVSNANOTVNLTVLQOPFERWQKEVGLKELIHQISAFISQSGSL 1380
 QY 1381 GKQONVGVKLYGRK 1394
 DB 1381 GKQONVGVKLYGRK 1394

DB 1381 GKQONVGVKLYGRK 1394
 RESULT 4
 PCT-US03-05226-7
 Sequence 7, Application EC/TUS0305226
 GENERAL INFORMATION:
 APPLICANT: St. Gene, Joseph W.
 TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 FILE REFERENCE: A-59941-1/RT/DCF/DHR
 CURRENT APPLICATION NUMBER: PCT/US03/05226
 CURRENT FILING DATE: 2003-02-18
 PRIOR APPLICATION NUMBER: US 09/296,731
 PRIOR FILING DATE: 1994-10-25
 PRIOR APPLICATION NUMBER: US 09/839,996
 PRIOR FILING DATE: 2001-04-20
 NUMBER OF SEQ ID NOS: 58
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 7
 LENGTH: 1395
 TYPE: PRT
 ORGANISM: Haemophilus influenzae
 PCT-US03-05226-7
 Query Match 99.6%; Score 7246.5; DB 1; Length 1395;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1393; Conservative 0; Mismatches 0; Indels 3; Gaps 2;
 QY 1 MKKTVFRLNPLFACISGIVSOAMAGHTYFGIDYQYRDFAEKNGKFTVQKNIKYNNQ 60
 DB 1 MKKTVFRLNPLFACISGIVSOAMAGHTYFGIDYQYRDFAEKNGKFTVQKNIKYNNQ 60
 QY 61 GQIVGSMWKAPMIDSVSRNGVAALVENQYTVSAHVNGYTTVDGABGNPDDHRT 120
 DB 61 GQIVGSMWKAPMIDSVSRNGVAALVENQYTVSAHVNGYTTVDGABGNPDDHRT 120
 QY 121 YKIVKRNKYKKNLHPEDDYENPLHFTYEAAPIDMTSNNGSYSDRTKPERVRIG 180
 DB 121 YKIVKRNKYKKNLHPEDDYENPLHFTYEAAPIDMTSNNGSYSDRTKPERVRIG 180
 QY 181 SGROFMRDQDQVAGAVHYLTAGNTHNRQAGNGSYLGGDVRAAGEYGLPIAGSK 240
 DB 181 SGROFMRDQDQVAGAVHYLTAGNTHNRQAGNGSYLGGDVRAAGEYGLPIAGSK 240
 QY 241 GDSGSMFTYDAEKQKMLNGILRGNPFGKENGQVYKSYFPELPERDLTSLYTA 300
 DB 241 GDSGSMFTYDAEKQKMLNGILRGNPFGKENGQVYKSYFPELPERDLTSLYTA 300
 QY 301 GNGSVTTISGNDGQSLTOKSGIPSEKITLANMSPLKXKDVHNPIDGNITSPRLN 360
 DB 301 GNGSVTTISGNDGQSLTOKSGIPSEKITLANMSPLKXKDVHNPIDGNITSPRLN 360
 QY 361 NGELTYPMQKQKSLIFASDINOGAGLYFEENFTVPSNSNOTMOGAGIHVSNSITVTK 420
 DB 361 NGELTYPMQKQKSLIFASDINOGAGLYFEENFTVPSNSNOTMOGAGIHVSNSITVTK 420
 QY 421 VNGVEHDLRSKIGKGLTHVQAKGENKGSISVGDGKVLLEQADDOGNQAFSEIGVSGR 480
 DB 421 VNGVEHDLRSKIGKGLTHVQAKGENKGSISVGDGKVLLEQADDOGNQAFSEIGVSGR 480
 QY 481 GTVQANDKQPTDKFYFGGRGLDNGHSLTFKRIQNTDEGAMIVHNTTOAANVTIT 540
 DB 481 GTVQANDKQPTDKFYFGGRGLDNGHSLTFKRIQNTDEGAMIVHNTTOAANVTIT 540
 QY 541 GNESTIVLPNGNNINKLDYRKELIANGMFGETDKNKNRNLILYKTTEDRTLLSGGN 600
 DB 541 GNESTIVLPNGNNINKLDYRKELIANGMFGETDKNKNRNLILYKTTEDRTLLSGGN 600
 QY 601 LKGDITQTKGLFESGRTPHAVNHLNKSSEMGIPQGETIWDHDMINRTFKAEFQIK 660
 DB 601 LKGDITQTKGLFESGRTPHAVNHLNKSSEMGIPQGETIWDHDMINRTFKAEFQIK 660

QY	661	GGAAVAVSNVSI	IGAMVTYSNNAATGAVWNOVTI	CTSDW	TGLTTCQAVD	TTTKYI	720		
Dp	661	GGAAVAVSNVSI	IEGMATYSNNAATGAVWNOVTI	CTSDW	TGLTTCQAVD	TTTKYI	720		
QY	661	GGAAVAVSNVSI	IEGMATYSNNAATGAVWNOVTI	CTSDW	TGLTTCQAVD	TTTKYI	720		
Dp	661	GGAAVAVSNVSI	IEGMATYSNNAATGAVWNOVTI	CTSDW	TGLTTCQAVD	TTTKYI	720		
QY	721	NSIPKTOINGS	INTDANATANYKGLAKINGNTLT	INHSQFT	LSNNA	TOIGNIRLSDNSTA	780		
Dp	721	NSIPKTOINGS	INTDANATANYKGLAKINGNTLT	INHSQFT	LSNNA	TOIGNIRLSDNSTA	780		
QY	781	TYNANLNGNTH	LDSAQPSLKNKSPHIOIQDKGT	TYLLENK	TWMPESD	TLTONTLNN	840		
Dp	781	TYNANLNGNTH	LDSAQPSLKNKSPHIOIQDKGT	TYLLENK	TWMPESD	TLTONTLNN	840		
QY	841	STTTTNSAVS	ASNSNTP - -RRRSLETET	TTSAERFNTLV	NGKLSGGXT	FOFTSLFG	898		
Dp	841	STTTTNSAVS	ASNSNTPRRRRSLETET	TTSAERFNTLV	NGKLSGGXT	FOFTSLFG	900		
QY	899	YKSDKALSNAB	EDTLYLSTRV	TGKEPE	TLTBO	LTVESKNQPLSNK	KTLTENDVDG	958	
Dp	901	YKSDKALSNAB	EDTLYLSTRV	TGKEPE	TLTBO	LTVESKNQPLSNK	KTLTENDVDG	960	
QY	959	ALSKYLVKND	DEFILHNP	IKEDLHND	LVBAEQARTLEAC	YVEPAKT	GTGEPKYSRR	101.88	
Dp	961	ALSKYLVKND	DEFILHNP	IKEDLHND	LVBAEQARTLEAC	YVEPAKT	GTGEPKYSRR	102.00	
QY	1019	ABBAFPPTL	PDOSL	NALBAKQAL	TAF	TOKSKAKTKYR	SKRAVFSDDLDS	107.98	
Dp	1021	ABBAFPPTL	PDOSL	NALBAKQAL	TAF	TOKSKAKTKYR	SKRAVFSDDLDS	108.00	
QY	1079	AALEVIDA	QOSECDRLAQB	EAKOROKO	LISRYNSAL	SEL	SAVNSML	STVODELRL	113.88
Dp	1081	AALEVIDA	QOSECDRLAQB	EAKOROKO	LISRYNSAL	SEL	SAVNSML	STVODELRL	114.00
QY	1139	FVDOASAV	WTNIAHODRRYDS	DAFRAYQ	QOKTN	RLQ	GVOKALANG	RLGAVFSHSKSN	113.98
Dp	1141	FVDOASAV	WTNIAHODRRYDS	DAFRAY	QOKTN	RLQ	GVOKALANG	RLGAVFSHSKSN	113.99
QY	1199	TFPEOVRNAT	LTWMSQFQXQ	QGLQ	QGVNCTG	SASWMAE	QSRKHKA	INVSVA	125.98
Dp	1200	TFPEOVRNAT	LTWMSQFQXQ	QGLQ	QGVNCTG	SASWMAE	QSRKHKA	INVSVA	125.99
QY	1259	SYOFRLQ	GLQIPFEGVNR	YTERENTQ	QSEVAV	KTPSLA	ENRYNG	ISVDTPTDNI	131.88
Dp	1260	SYOFRLQ	GLQIPFEGVNR	YTERENTQ	QSEVAV	KTPSLA	ENRYNG	ISVDTPTDNI	131.89
QY	1319	SVKPYFVNT	VDVSNANVQ	TNLT	LTLOQ	PFGRWQ	KVGSJKA	ILHFOISAF	137.88
Dp	1320	SVKPYFVNT	VDVSNANVQ	TNLT	LTLOQ	PFGRWQ	KVGSJKA	ILHFOISAF	137.89
QY	1379	QLGKQKQ	VGWKLGYR	1394					
Dp	1380	QLGKQKQ	VGWKLGYR	1395					
RESULT 5									
US-10-687-046-7									
Sequence 7, Application US/10687046									
GENERAL INFORMATION:									
APPLICANT: St. Gene, Joseph W.									
TITLE OF INVENTION: HAZMOPIHUS ADHERENCE AND PENETRATION PROTEINS									
FILE REFERENCE: A-59941-1/REF/DC/DDR									
CURRENT APPLICATION NUMBER: US/10/687,046									
CURRENT FILING DATE: 2003-10-15									
PRIOR APPLICATION NUMBER: US/10/080,505									
PRIOR FILING DATE: 2003-02-22									
PRIOR APPLICATION NUMBER: US 08/296,791									
PRIOR FILING DATE: 1994-10-25									
PRIOR APPLICATION NUMBER: US 09/639,996									
PRIOR FILING DATE: 2001-04-20									

Query Match	99.6%	Score 7246.5	DB 31	Length 1395
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1393	Conservative 0	Mismatches 0	Indels 3	Gaps 2
Query	1	MKTYVFLNLTACISLIGVQAWAGTGYGIDQYRDFAEKGRFTWGAONIKYTKQ	60	
Db	1	MKTYVFLNLTACISLIGVQAWAGTGYGIDQYRDFAEKGRFTWGAONIKYTKQ	60	
Qy	6	GOLVGTSMTAPMIDFVSYSVRNGVAALAYENQYIVSAANVGYTDVDFGAEKNPDQHFPT	120	
Db	6	GOLVGTSMTAPMIDFVSYSVRNGVAALAYENQYIVSAANVGYTDVDFGAEKNPDQHFPT	120	
Qy	12	YKIVKRNKYKQNLAPYEDDYHNPRLHKFVTEAPIDMTSMNGSTYSDBTKYPERVIG	180	
Db	12	YKIVKRNKYKQNLAPYEDDYHNPRLHKFVTEAPIDMTSMNGSTYSDBTKYPERVIG	180	
Qy	18	SGRQFPMNDQKEDQYAGAGHYTLTRGNTNHOAGAAGVSYLGEQVRAAGYGPLPLASK	240	
Db	18	SGRQFPMNDQKEDQYAGAGHYTLTRGNTNHOAGAAGVSYLGEQVRAAGYGPLPLASK	240	
Qy	24	GDGSPHEITDAKQKWLINGILREGNPEFGEKNGFOLVRSKSYDELFERDLHTSLYTRA	300	
Db	24	GDGSPHEITDAKQKWLINGILREGNPEFGEKNGFOLVRSKSYDELFERDLHTSLYTRA	300	
Qy	30	GNQVYTTISGNDNGQSSITQKSGIPSEIKITLANMSLPLAKOKYHNPRYDGNLYSPRLN	360	
Db	30	GNQVYTTISGNDNGQSSITQKSGIPSEIKITLANMSLPLAKOKYHNPRYDGNLYSPRLN	360	
Qy	36	NGEFLTFMDQKQSLIFASDINQAGGLYEGNFTVSNSSQCMQAGHIVSENSLVTK	420	
Db	36	NGEFLTFMDQKQSLIFASDINQAGGLYEGNFTVSNSSQCMQAGHIVSENSLVTK	420	
Qy	42	VNGVEHRLSKIKGTGTHVQAKGENKKSIVSEBQSKVILEQADQGNKQAFSEIGVSGR	480	
Db	42	VNGVEHRLSKIKGTGTHVQAKGENKKSIVSEBQSKVILEQADQGNKQAFSEIGVSGR	480	
Qy	48	GTVOIANDKQFDTDKEYFGFRQRLDINSHSLTFKRIQNTDEGAMIVNHTTOANVTIT	540	
Db	48	GTVOIANDKQFDTDKEYFGFRQRLDINSHSLTFKRIQNTDEGAMIVNHTTOANVTIT	540	
Qy	54	GNESIVLPNGNNINMLDPRKELANNGRFGTDQNGNRLMYKPTERLTLISGGN	600	
Db	54	GNESIVLPNGNNINMLDPRKELANNGRFGTDQNGNRLMYKPTERLTLISGGN	600	
Qy	60	LNGDITOTKGLPFSGRPLPHAVYHLKRRSEHEGIPQELVMDHWNINTEFAEKFPOLK	660	
Db	60	LNGDITOTKGLPFSGRPLPHAVYHLKRRSEHEGIPQELVMDHWNINTEFAEKFPOLK	660	
Qy	66	GGSAVSVNYSIRKMTVSNANATGCVVPOQNTITCRSDMTGLTTCQVDLDTXYI	720	
Db	66	GGSAVSVNYSIRKMTVSNANATGCVVPOQNTITCRSDMTGLTTCQVDLDTXYI	720	
Qy	72	NSIPETQJNSINULTNNAIYKGLACNNNTLTNHSOFTLSNNAQIGNIRLSDNSTA	780	
Db	72	NSIPETQJNSINULTNNAIYKGLACNNNTLTNHSOFTLSNNAQIGNIRLSDNSTA	780	
Qy	78	TYDANANNGVHLTDAQSLKSHSFHQIQGDKGTVYLENATWTPSDTTLQNTLNN	840	
Db	78	TYDANANNGVHLTDAQSLKSHSFHQIQGDKGTVYLENATWTPSDTTLQNTLNN	840	
Qy	84	STITLNSAYSASSNTP--RRRSLETTPTSAHRFNTLTVNGKLSGGSTFOFTSSLFG	898	
Db	84	STITLNSAYSASSNTP--RRRSLETTPTSAHRFNTLTVNGKLSGGSTFOFTSSLFG	898	
Qy	89	YSDKGLKISDADEYTIISVNTGKEPPTLEQITLVESQNDQPSDGLKTYLLENHVNAG	958	
Db	90	YSDKGLKISDADEYTIISVNTGKEPPTLEQITLVESQNDQPSDGLKTYLLENHVNAG	960	
Qy	95	ARRYLVGNQGBERLNP--KQQLHLNLTVAQAEBVTELEAQVEPTAKTGTGEPKYSRR	1018	

```

Db 96: AATKLVKNDGEFRLNPLIKQELHNDLVRADQAEFTLEAKQVEPTAKTQTEGPKHSR 1020
Qy 1029 AARAFEDTLPDQSLINLLEAKQELTLETOKSKAKTKQVSKRAVEFSDPLDQSLFALB 1078
Db 1032 AARAFEDTLPDQSLINLLEAKQELTLETOKSKAKTKQVSKRAVEFSDPLDQSLFALB 1080
Qy 1079 AALEYINAPQOSKCKLAGEAEKQKQKQDLISRSNSALSELATYNSMLSVODELDEL 1138
Db 1081 AALEYIDAPQOSKCKLAGEAEKQKQKQDLISRSNSALSELATYNSMLSVODELDEL 1140
Qy 1139 FVDAQASAVWCTNADQKRYDSDAFRAYQQQKTNLRQIGVOKALANGRIGAVFHSRSDN 1198
Db 1141 FVDAQASAVWCTNADQKRYDSDAFRAYQQQKTNLRQIGVOKALANGRIGAVFHSRSDN 1199
Qy 1199 FVDAQASAVWCTNADQKRYDSDAFRAYQQQKTNLRQIGVOKALANGRIGAVFHSRSDN 1258
Db 1200 FVDAQASAVWCTNADQKRYDSDAFRAYQQQKTNLRQIGVOKALANGRIGAVFHSRSDN 1259
Qy 1259 SVQRLGQIGIOPRYGNRYFIERENYQSEBVRKTPSLAENRYNGIENVDTPTDNI 1318
Db 1260 SVQRLGQIGIOPRYGNRYFIERENYQSEBVRKTPSLAENRYNGIENVDTPTDNI 1319
Qy 1319 SVKEFFVNVYDVNSANVQTTNLTVLQCPFGRYQKQVGLKAEILHFOISATFSKQGS 1378
Db 1320 SVKEFFVNVYDVNSANVQTTNLTVLQCPFGRYQKQVGLKAEILHFOISATFSKQGS 1379
Qy 1379 QLGKQNVGVKLGTYM 1394
Db 1380 QLGKQNVGVKLGTYM 1395

RESULT 6
PCT-US03-05226-9
; Sequence 9, Application PC/TUS0305226
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RT/DCP/DHR
; CURRENT FILING DATE: 2003-02-18
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1434
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (568)..(568)
; OTHER INFORMATION: The 'Xaa' at location 568 stands for Ser, Gly, Arg, or Cys.
; NAME/KEY: misc_feature
; LOCATION: (1702)..(1702)
; OTHER INFORMATION: "n" at position 1702 can be any base.
PCT-US03-05226-9

Query Match: 84.1%; Score 6115; DB 1; Length 1434;
Best Local Similarity: 82.9%; Pred. No. 0;
Matches 1194; Conservative 81; Mismatches 113; Indels 52; Gaps 10;

```

```

Db 121 YKIVKNNYNDADHGRYLDYNNPRLHFKVITDAAPIDMTSHDQKAKTAKKEPPEYKVG 180
Qy 181 SGRQFNRDQKXPQVAGAYHYLTAGNTHQVQAGNGSYSLGQVYKXGEXGLPIAGSK 240
Db 181 SGRQVWDDQNNRNTYSLDGYNTLTGANTYNSGSGGQGYVANGIRKXGQVGLPIASSF 240
Qy 241 GDSGPMITTYAEKQKMLINGILBEGNPFQKENGSQLYKSYFPELPERDHTLSLYTA 300
Db 241 GDSGPMITTYAEKQKMLINGILBEGNPFQKENGSQLYKSYFPELPERDHTLSLYTA 300
Qy 301 GNGVITISGNNQKGSITQKSGIPSEIKITLANNSELEKQKYNHNPYDQGPNTYPRLN 360
Db 301 GNGVITISGNNQKGSITQKSGIPSEIKITLANNSELEKQKYNHNPYDQGPNTYPRLN 360
Qy 361 NSELTIFMDQKQSLIFASPDINQAGAGLYFEGNFTVSPNSNQTQWQAGIHVSESTYTK 420
Db 359 NSELTIFMDQKQSLIFASPDINQAGAGLYFEGNFTVSPNSNQTQWQAGIHVSESTYTK 418
Qy 421 VNGTEHRLSTKIGKTLHYQAKENGKGSISYVQKTYLLEQVQDDQKQAFSEIGVSR 480
Db 419 VNGTEHRLSTKIGKTLHYQAKENGKGSISYVQKTYLLEQVQDDQKQAFSEIGVSR 478
Qy 481 GTVQANDKQPDQKPYFGFGGRLDIAKESLTPKRIQNTDEGAMIVNHTTQAAVNTIT 540
Db 479 GTVQANDKQPDQKPYFGFGGRLDIAKESLTPKRIQNTDEGAMIVNHTTQAAVNTIT 538
Qy 541 GNESTIPLP-NGNNINKLDYRKSIAYVNGFEGTDKNNKAKMLIYKPTTEDRTLLSGGT 599
Db 539 GNESTIPLP-NGNNINKLDYRKSIAYVNGFEGTDKNNKAKMLIYKPTTEDRTLLSGGT 598
Qy 600 NLKQDITQTKGLFSGRPPTPRAYNHLNKRSEMEGIPQGLIWMDDMINTPRANFQI 659
Db 599 NLKQDITQTKGLFSGRPPTPRAYNHLNKRSEMEGIPQGLIWMDDMINTPRANFQI 658
Qy 660 KGSAAVSNVSYIEGKMTVSNANATFGVYVQWQNTICRSDWGLTTCQKVDLDTYV 719
Db 659 KGSAAVSNVSYIEGKMTVSNANATFGVYVQWQNTICRSDWGLTTCQKVDLDTYV 718
Qy 720 INSIPATQINSINLTNATNATVANKGLAKNGVTLNHSOFTLSNNAFOINRILS---- 775
Db 719 INSIPATQINSINLTNATNATVANKGLAKNGVTLNHSOFTLSNNAFOINRILS---- 778
Qy 776 ----DNST-----ATVDNANLNGVHLTDSKPS 800
Db 779 ARVNNATLMGDVNLADTSFTLSNQAQIGTISLHQDQATVDNANLNGVHLTDSKPS 838
Qy 801 LKNSHPSHQIQDKGIVTLLENATVWPSDPTLQNTLNNSTTLNSAYSAASNTP-RR 859
Db 839 LKNSHPSHQIQDKGIVTLLENATVWPSDPTLQNTLNNSTTLNSAYSAASNTP-RR 898
Qy 860 RSLFETTPYSAERFVTLTVNGKLSGGTQFTSSIFGYKSDKLSNDEGTYLTVR 919
Db 899 RSLFETTPYSAERFVTLTVNGKLSGGTQFTSSIFGYKSDKLSNDEGTYLTVR 958
Qy 920 NTGKEPTELEUTLVESKQNPISDKLKEFTLENDHVDAGLERYKLVKNDGEFRLNPLIK 979
Db 959 NTGKEPTELEUTLVESKQNPISDKLKEFTLENDHVDAGLERYKLVKNDGEFRLNPLIK 1018
Qy 980 QELHNDLVRADQAEFTLEAKQVEPTAKTQTEGPKHSRRAAPFDTLPQSLINLLEA 1039
Db 1019 QELHNDLVRADQAEFTLEAKQVEPTAKTQTEGPKHSRRAAPFDTLPQSLINLLEA 1075
Qy 1040 KQELTLETOKSKAKTKQVSKRAVEFSDPLDQSLFALB 1095
Db 1076 KQELTLETOKSKAKTKQVSKRAVEFSDPLDQSLFALB 1135
Qy 1096 AQ-EEAEKQKQKQDLISRSNSALSELATYNSMLSVODELDELVDQASAVWCTNAD 1154
Db 1136 AQEEAEKQKQKQDLISRSNSALSELATYNSMLSVODELDELVDQASAVWCTNAD 1195
Qy 1155 KRYSDAFRAYQQQKTNLRQIGVOKALANGRIGAVFHSRSDNFTDQVYKNNATLTNMS 1214

```

```

Db      1196 KRRYSDAFRAY-QQKTNLRQIGVOKALDNCRIGAVFSHSNDTFDEQYKNAHTLTWMS 1254
Qy      1215 GFAOYQWMDLQPGVNVGTGISASRNAEQQSKTHRRALNYGNASVQFRLGQIGIOPYFG 1274
Db      1255 GFAOYQWMDLQPGVNVGTGISASRNAEQQSKTHRRALNYGNASVQFRLGQIGIOPYFG 1314
Qy      1275 VNRYPFERENYQSEEVAVKTPSLAFRRYNAGIRVDYTFPTDNIISVKYFFVNVVNSNA 1334
Db      1315 VNRYPFERENYQSEEVAVKTPSLAFRRYNAGIRVDYTFPTDNIISVKYFFVNVVNSNA 1374
Qy      1335 NVQTTVNTLVYQOPFGVYQKQEVGKAEILHFOISAFISKQSQSLQKQKQVGVKLGYSR 1394
Db      1375 NVQTTVNTLVYQOPFGVYQKQEVGKAEILHFOISAFISKQSQSLQKQKQVGVKLGYSR 1434

RESULT 7
US-10-687-046-9
; Sequence 9, Application US/10687c46
; GENERAL INFORMATION:
; APPLICANT: St. Geme, Joseph M.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
; CURRENT APPLICATION NUMBER: US/10/687,046
; CURRENT FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: US/10/080,505
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-12-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SORTAB: Patent version 3.1
; SEQ ID NO 9
; LENGTH: 1434
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (368)..(568)
; OTHER INFORMATION: The 'Xaa' at location 568 stands for Ser, Gly, Arg, or Cys.
; NAME/KEY: misc_feature
; LOCATION: (1702)..(1702)
; OTHER INFORMATION: "n" at position 1702 can be any base.
US-10-687-046-9

Query Match      84.1%; Score 6115; DB 31; Length 1434;
Best Local Similarity 82.9%; Pred. No. 0;
Matches 1194; Conservative 81; Mismatches 113; Indels 52; Gaps 10;

Qy      1 MKTVELNLFTACISLGIVSQAAGHTYFGIDYQYRDPDAENKKGKFTYGAQNTIKYNNQ 60
Db      1 MKTVELNLFTACISLGIVSQAAGHTYFGIDYQYRDPDAENKKGKFTYGAQNTIKYNNQ 60
Qy      61 GQLVGTSMTKAMPIDFSVWSKGVYALVGGQYIVSAHNVGICNVDFGAEQNDQRFET 120
Db      61 GQLVGTSMTKAMPIDFSVWSKGVYALVGGQYIVSAHNVGICNVDFGAEQNDQRFET 120
Qy      121 YKIVKNNYKKNLHPYEDYHNPRILHKEVTEAIPIDMTSNNGSTYSRRTKYEPRVRLG 180
Db      121 YKIVKNNYKKNLHPYEDYHNPRILHKEVTEAIPIDMTSNNGSTYSRRTKYEPRVRLG 180
Qy      181 SGQFWENDQDKGDVYAGAHYILTAQNTNORGAGNKSYLGQDVAKAGVGPPIPLNGSK 240
Db      181 SGQFWENDQDKGDVYAGAHYILTAQNTNORGAGNKSYLGQDVAKAGVGPPIPLNGSK 240
Qy      241 GDSGSPMETYDAEKQKLLINGILBENPREGNGFQLYVRSYFDEIFPRDLHTSLTSA 300
Db      241 GDSGSPMETYDAEKQKLLINGILBENPREGNGFQLYVRSYFDEIFPRDLHTSLTSA 300
Qy      301 GNGVYITISGNDGQGSITQKSGIPSEIKITLANNSLPLKSKDXYVHNPRYDGNITYSBR 360
Db      301 GNGVYITISGNDGQGSITQKSGIPSEIKITLANNSLPLKSKDXYVHNPRYDGNITYSBR 360

```

```

Qy      361 NGELTYPMQKQGSLLIPASDINQAGLYFEGNFTVSPNSNMQMGAGHSENVSTYWK 420
Db      369 NGKNIYFGEELGSLTLTDTDDQAGGLYFBDFTVSTKNEWTKAGAHVSEISTYWK 418
Qy      421 VNGVEHRLSKIKGKTLHVAQKNGKSGISVGDGVILIEQADQGNKQAFSEIGLYSGR 480
Db      419 VNGVEHRLSKIKGKTLHVAQKNGKSGISVGDGVILIEQADQGNKQAFSEIGLYSGR 478
Qy      481 GTVQLNDKQPTDKEFYFPRGRLDNLNHSILFRLQNTDEGAMVYNNHTQAANTIT 540
Db      479 GTVQLNDKQPTDKEFYFPRGRLDNLNHSILFRLQNTDEGAMVYNNHTQAANTIT 538
Qy      541 GNESTVLP--NGNNYKLDYRKEILA YNGVPEGDYKKNHGRNLNLYPTEDEDTLLSGGT 599
Db      539 GNESTVLP--NGNNYKLDYRKEILA YNGVPEGDYKKNHGRNLNLYPTEDEDTLLSGGT 598
Qy      600 NLKGDITQTKGLFESGSRPTPHAVNHLNKKWSEMEGIPQGEI VWDHDMINRTPKAEFOI 659
Db      599 NLKGDITQTKGLFESGSRPTPHAVNHLNKKWSEMEGIPQGEI VWDHDMINRTPKAEFOI 658
Qy      660 KQGSAYVSRVSSIEGNTVSNANATFEGVYVNOONTICTSDPTGLTTCQKVDLTDYK 719
Db      659 KQGSAYVSRVSSIEGNTVSNANATFEGVYVNOONTICTSDPTGLTTCQKVDLTDYK 718
Qy      720 INSI PKQINGSLNLTDAATANVGLAKLNGANTLTNHSOFTLSNNTQIGNIRLS--- 775
Db      719 IDSIPTQINGSLNLTDAATANVGLAKLNGANTLTNHSOFTLSNNTQIGNIRLS--- 778
Qy      776 ---DNST-----ATVDNANLNGVHLTSDAQS 800
Db      779 AAVNNATLMQDVVLADTSFTLSNQATQIGTISLHQQAQAVDANLNGVHLTSDAQS 838
Qy      801 LNSHFSHQIQSGKGTVTLENATWYPSDTLLQNTLANSITTLNSAYASANNTP--RR 859
Db      839 LNSHFSHQIQSGKGTVTLENATWYPSDTLLQNTLANSITTLNSAYASANNTP--RR 898
Qy      860 RSLFETTPYSAHRNTLTNNGKLSGGGTFQCTSSLGYSYDGLKLSNABEDYTLSTR 919
Db      899 RSLFETTPYSAHRNTLTNNGKLSGGGTFQCTSSLGYSYDGLKLSNABEDYTLSTR 958
Qy      920 NNGKEPTELBQTLTVESKDNQPLSDPKLFTLENDHVDAGALRYKLYVNDSEFLNHPKE 979
Db      959 NNGKEPTELBQTLTVESKDNQPLSDPKLFTLENDHVDAGALRYKLYVNDSEFLNHPKE 1018
Qy      980 QELNNDIVRAQAEKRLTBAKQVEPTATOTGEPKVSRRMAAARPTTLDDQSLMALEA 1039
Db      1019 QELNNDIVRAQAEKRLTBAKQVEPTATOTGEPKVSRRMAAARPTTLDDQSLMALEA 1075
Qy      1040 KQAEITAEQKRAKTKKYSKRAV---FSDPLDQS-LFALAEALVYDAPQSEKRL 1095
Db      1076 KQAEITAEQKRAKTKKYSKRAV---FSDPLDQS-LFALAEALVYDAPQSEKRL 1135
Qy      1096 AQ-EAEKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1154
Db      1136 AQ-EAEKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 1195
Qy      1155 KRRYSDAFRAYQQQKTNLRQIGVOKALDNCRIGAVFSHSNDTFDEQYKNAHTLTWMS 1214
Db      1196 KRRYSDAFRAY-QQKTNLRQIGVOKALDNCRIGAVFSHSNDTFDEQYKNAHTLTWMS 1254
Qy      1215 GFAOYQWMDLQPGVNVGTGISASRNAEQQSKTHRRALNYGNASVQFRLGQIGIOPYFG 1274
Db      1255 GFAOYQWMDLQPGVNVGTGISASRNAEQQSKTHRRALNYGNASVQFRLGQIGIOPYFG 1314
Qy      1275 VNRYPFERENYQSEEVAVKTPSLAFRRYNAGIRVDYTFPTDNIISVKYFFVNVVNSNA 1334
Db      1315 VNRYPFERENYQSEEVAVKTPSLAFRRYNAGIRVDYTFPTDNIISVKYFFVNVVNSNA 1374
Qy      1335 NVQTTVNTLVYQOPFGVYQKQEVGKAEILHFOISAFISKQSQSLQKQKQVGVKLGYSR 1394
Db      1375 NVQTTVNTLVYQOPFGVYQKQEVGKAEILHFOISAFISKQSQSLQKQKQVGVKLGYSR 1434

```

```

RESULT 3
PCT-US03-05226-17
; Sequence 17, Application PC/TUS0305226
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
; CURRENT APPLICATION NUMBER: PCT/US03/05226
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1411
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
PCT-US03-05226-17

```

Query Match 80.1%; Score 5827.5; DB 1; Length 1411;

Best local similarity 79.5%; Ered. No. 0;

Matches 1133; Conservative 107; Mismatches 139; Indels 47; Gaps 12;

```

QY 1 MKTVPRLNLTACISLGSIVQAWAGHTYFGIDYQYDEAKKGFVGAONIKVYKQ 60
DB 1 MKTVPRLNLTACISLGSIVQAWAGHTYFGIDYQYDEAKKGFVGAONIKVYKQ 60
QY 61 GOLVSTMTAPMIDPSVSRNGVALYENQYIVANVGYTDVDFGAGNPNPDRFT 120
DB 61 GNLVSTMTAPMIDPSVSRNGVALYENQYIVANVGYTDVDFGAGNPNPDRFT 120
QY 121 YKIVKNNYKQCNAPYEDDYHNPLHKFTPAEIDMTSMNASTYSDRTKCPERYRG 180
DB 121 YKIVKNNYKQCNAPYEDDYHNPLHKFTPAEIDMTSMNASTYSDRTKCPERYRG 180
QY 122 YKIVKNNYKQCNAPYEDDYHNPLHKFTPAEIDMTSMNASTYSDRTKCPERYRG 180
DB 122 YKIVKNNYKQCNAPYEDDYHNPLHKFTPAEIDMTSMNASTYSDRTKCPERYRG 180
QY 181 SGROPWRNDQKGDVAGAYHYLTAGNTNORAGAGNSYVSGDVRKAGEYEPFLAGSK 240
DB 181 SGROPWRNDQKGDVAGAYHYLTAGNTNORAGAGNSYVSGDVRKAGEYEPFLAGSK 240
QY 241 GDSGPMETIDAEKQWILNGILRENGPEFGKNGFOLVRASYDRIPEKLIHSITRA 300
DB 241 GDSGPMETIDAEKQWILNGILRENGPEFGKNGFOLVRASYDRIPEKLIHSITRA 300
QY 301 GNGVYITSGNDNGSGSITQKSGIPSEIKITLANMSLPLKEDKXVHPRYDGNITSPRLN 360
DB 301 GNGVYITSGNDNGSGSITQKSGIPSEIKITLANMSLPLKEDKXVHPRYDGNITSPRLN 360
QY 301 GNGVYITSGNDNGSGSITQKSGIPSEIKITLANMSLPLKEDKXVHPRYDGNITSPRLN 360
DB 301 GNGVYITSGNDNGSGSITQKSGIPSEIKITLANMSLPLKEDKXVHPRYDGNITSPRLN 360
QY 361 NGETLYENQKQSLIFASBINOGAGSLYFEGNFTVSNSTQTMQAGIHVSENSTYTWK 420
DB 361 NGETLYENQKQSLIFASBINOGAGSLYFEGNFTVSNSTQTMQAGIHVSENSTYTWK 420
QY 361 NGETLYENQKQSLIFASBINOGAGSLYFEGNFTVSNSTQTMQAGIHVSENSTYTWK 420
DB 361 NGETLYENQKQSLIFASBINOGAGSLYFEGNFTVSNSTQTMQAGIHVSENSTYTWK 420
QY 421 VNGVHDLSTKIGKGLHVKAGKNGKGSISVGGVAYILEGADQGNKQAPSEIGVSR 480
DB 421 VNGVHDLSTKIGKGLHVKAGKNGKGSISVGGVAYILEGADQGNKQAPSEIGVSR 480
QY 421 VNGVHDLSTKIGKGLHVKAGKNGKGSISVGGVAYILEGADQGNKQAPSEIGVSR 480
DB 421 VNGVHDLSTKIGKGLHVKAGKNGKGSISVGGVAYILEGADQGNKQAPSEIGVSR 480
QY 481 GTVQVNDKQKFTKXKYPFGRLDLNGHSLFTFKRIQNTDEGAMVNNNTQAAVYIT 540
DB 481 GTVQVNDKQKFTKXKYPFGRLDLNGHSLFTFKRIQNTDEGAMVNNNTQAAVYIT 540
QY 479 GTVQVNDKQKFTKXKYPFGRLDLNGHSLFTFKRIQNTDEGAMVNNNTQAAVYIT 540
DB 479 GTVQVNDKQKFTKXKYPFGRLDLNGHSLFTFKRIQNTDEGAMVNNNTQAAVYIT 540
QY 541 GNESTVLPNGNNINLQDYKELAYNGFGRTDKANKANRLNLYKPTEDRILLISGGIN 600
DB 541 GNESTVLPNGNNINLQDYKELAYNGFGRTDKANKANRLNLYKPTEDRILLISGGIN 600
QY 539 GSDYI---NDNTGDIITNKRDIAPNGVFGKDDTKNTRNLVNTPLNDKHFLLISGGIN 594
DB 539 GSDYI---NDNTGDIITNKRDIAPNGVFGKDDTKNTRNLVNTPLNDKHFLLISGGIN 594
QY 601 LKGDITQCKGLFFSGRPTPHAYNHLKPKSEWEGIPQGEIWMDDHINRTKAKENFOIK 660
DB 601 LKGDITQCKGLFFSGRPTPHAYNHLKPKSEWEGIPQGEIWMDDHINRTKAKENFOIK 660
QY 595 LKGNITQDGGTIVFSRPTPHAYNHLKPKSEWEGIPQGEIWMDDHINRTKAKENFOIK 652
DB 595 LKGNITQDGGTIVFSRPTPHAYNHLKPKSEWEGIPQGEIWMDDHINRTKAKENFOIK 652
QY 661 GGSAYVSRNVSIEGNMTVSNANATGVVWQONTICTRSDMTGLTTCQKDLITTKVI 720
DB 661 GGSAYVSRNVSIEGNMTVSNANATGVVWQONTICTRSDMTGLTTCQKDLITTKVI 720
QY 653 GGSAYVSRNVSIEGNMTVSNANATGVVWQONTICTRSDMTGLTTCQKDLITTKVI 712
DB 653 GGSAYVSRNVSIEGNMTVSNANATGVVWQONTICTRSDMTGLTTCQKDLITTKVI 712

```

```

QY 721 NSIPPTQINSINLTDMATPAVKGLAKUNGVTLTNHSQFTLSNNATQIGNILSPNSTA 780
DB 721 NSIPPTQINSINLTDMATPAVKGLAKUNGVTLTNHSQFTLSNNATQIGNILSPNSTA 780
QY 713 DSIPPTQINSINLTDMATPAVKGLAKUNGVTLTNHSQFTLSNNATQIGNILSPNSTA 772
DB 713 DSIPPTQINSINLTDMATPAVKGLAKUNGVTLTNHSQFTLSNNATQIGNILSPNSTA 772
QY 781 TVDNANLGNVHLDSKQSLKNSHFSHQISQDGTTLVNAWTWPSPTTQNTLNN 840
DB 781 TVDNANLGNVHLDSKQSLKNSHFSHQISQDGTTLVNAWTWPSPTTQNTLNN 840
QY 773 TVDNANLGNVHLDSKQSLKNSHFSHQISQDGTTLVNAWTWPSPTTQNTLNN 832
DB 773 TVDNANLGNVHLDSKQSLKNSHFSHQISQDGTTLVNAWTWPSPTTQNTLNN 832
QY 841 STTINSAYASASNTPR-RRLSETEPTPSAEHRENTLTUNGKLSGGQTFQFTSSLPFY 899
DB 841 STTINSAYASASNTPR-RRLSETEPTPSAEHRENTLTUNGKLSGGQTFQFTSSLPFY 899
QY 833 STTINSAYASASNTPR-RRLSETEPTPSAEHRENTLTUNGKLSGGQTFQFTSSLPFY 892
DB 833 STTINSAYASASNTPR-RRLSETEPTPSAEHRENTLTUNGKLSGGQTFQFTSSLPFY 892
QY 900 KSDKILSNDAGDYTLISVRNTGKPELTBLTLVSKNOQLSLKLTLENDHDAKA 959
DB 900 KSDKILSNDAGDYTLISVRNTGKPELTBLTLVSKNOQLSLKLTLENDHDAKA 959
QY 893 KSDKILSNDAGDYTLISVRNTGKPELTBLTLVSKNOQLSLKLTLENDHDAKA 952
DB 893 KSDKILSNDAGDYTLISVRNTGKPELTBLTLVSKNOQLSLKLTLENDHDAKA 952
QY 960 LRYKLVKNDGFERLNPDIKEQELHNDVYRAQAEFTLEKQVPEPAKTQTEGPEVRSRA 1019
DB 960 LRYKLVKNDGFERLNPDIKEQELHNDVYRAQAEFTLEKQVPEPAKTQTEGPEVRSRA 1019
QY 953 WRTLVKNGGFEFLNPFIEQELHNDVYRAQAEFTLEKQVPEPAKTQTEGPEVRSRA 1010
DB 953 WRTLVKNGGFEFLNPFIEQELHNDVYRAQAEFTLEKQVPEPAKTQTEGPEVRSRA 1010
QY 1020 ARAAPPDTLPDQSLNLEAKQAEFTLEKQVPEPAKTQTEGPEVRSRA 1064
DB 1020 ARAAPPDTLPDQSLNLEAKQAEFTLEKQVPEPAKTQTEGPEVRSRA 1064
QY 1011 -RAAFSDTPPQSQNLALQAEFTLEKQVPEPAKTQTEGPEVRSRA 1069
DB 1011 -RAAFSDTPPQSQNLALQAEFTLEKQVPEPAKTQTEGPEVRSRA 1069
QY 1065 FSDPDLDSL---FA-----LEALEVIDAQQSEKNDLAEPAKQKQKQD 1108
DB 1065 FSDPDLDSL---FA-----LEALEVIDAQQSEKNDLAEPAKQKQKQD 1108
QY 1070 LALVLEKSTQIDIPANAKLAQNLTLQLEKALAV---AEQKSKKQAEQAKQKQD 1126
DB 1070 LALVLEKSTQIDIPANAKLAQNLTLQLEKALAV---AEQKSKKQAEQAKQKQD 1126
QY 1109 LISRYNSALSELATVNSMLSVODELRLPVDQAQSAVNTNIAODKRYDSDAFRAYQ 1168
DB 1109 LISRYNSALSELATVNSMLSVODELRLPVDQAQSAVNTNIAODKRYDSDAFRAYQ 1168
QY 1127 LISRYNSALSELATVNSMLSVODELRLPVDQAQSAVNTNIAODKRYDSDAFRAYQ 1185
DB 1127 LISRYNSALSELATVNSMLSVODELRLPVDQAQSAVNTNIAODKRYDSDAFRAYQ 1185
QY 1169 OKTNIRQIGVQKALANGRIGAFVSHSRSDNTPEQVKNATLTMMSGRAYQKQDLOFV 1228
DB 1169 OKTNIRQIGVQKALANGRIGAFVSHSRSDNTPEQVKNATLTMMSGRAYQKQDLOFV 1228
QY 1186 OKTNIRQIGVQKALANGRIGAFVSHSRSDNTPEQVKNATLTMMSGRAYQKQDLOFV 1245
DB 1186 OKTNIRQIGVQKALANGRIGAFVSHSRSDNTPEQVKNATLTMMSGRAYQKQDLOFV 1245
QY 1229 NVGTGTSKSKAEKQSKHRAINYGNASVQFRLGOLGIPYGVNRYFEBENYQSE 1288
DB 1229 NVGTGTSKSKAEKQSKHRAINYGNASVQFRLGOLGIPYGVNRYFEBENYQSE 1288
QY 1246 NVGTGTSKSKAEKQSKHRAINYGNASVQFRLGOLGIPYGVNRYFEBENYQSE 1305
DB 1246 NVGTGTSKSKAEKQSKHRAINYGNASVQFRLGOLGIPYGVNRYFEBENYQSE 1305
QY 1289 EYVVKTPSLAENRINAGIRVDYFTPTDNISVKEYPFVNVYDVSNANQVTLNLYVLOP 1348
DB 1289 EYVVKTPSLAENRINAGIRVDYFTPTDNISVKEYPFVNVYDVSNANQVTLNLYVLOP 1348
QY 1306 EYVVKTPSLAENRINAGIRVDYFTPTDNISVKEYPFVNVYDVSNANQVTLNLYVLOP 1365
DB 1306 EYVVKTPSLAENRINAGIRVDYFTPTDNISVKEYPFVNVYDVSNANQVTLNLYVLOP 1365
QY 1349 EYVVKTPSLAENRINAGIRVDYFTPTDNISVKEYPFVNVYDVSNANQVTLNLYVLOP 1394
DB 1349 EYVVKTPSLAENRINAGIRVDYFTPTDNISVKEYPFVNVYDVSNANQVTLNLYVLOP 1394
QY 1366 EYVVKTPSLAENRINAGIRVDYFTPTDNISVKEYPFVNVYDVSNANQVTLNLYVLOP 1411
DB 1366 EYVVKTPSLAENRINAGIRVDYFTPTDNISVKEYPFVNVYDVSNANQVTLNLYVLOP 1411

```

```

RESULT 9
US-10-687-046-17
; Sequence 17, Application US/10687046
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
; CURRENT APPLICATION NUMBER: US/10/687,046
; CURRENT FILING DATE: 2003-10-15
; PRIOR APPLICATION NUMBER: US/10/080,505
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1411
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-687-046-17

```


Query Match 80.1%; Score 5827.5; DB 31; Length 1411;
 Best Local Similarity 79.5%; Pred. No. 0; Mismatches 139; Indels 47; Gaps 12;
 Matches 1133; Conservative 107; Mismatches 139; Indels 47; Gaps 12;

QY 1 MKKTVRLNPLFLACISLGIIVSQAAGHTYFGIDYQYRDPANRKGKFTVGAONIKYVNRQ 60
 DB 1 MKKTVRLNPLFLACISLGIIVSQAAGHTYFGIDYQYRDPANRKGKFTVGAONIKYVNRQ 60

QY 61 GQVGTSMTKAPMIDSVSVSRNGVAALYENQYIVSAHNVGTVDPGAGANNPDHRT 120
 DB 61 GQVGTSMTKAPMIDSVSVSRNGVAALYENQYIVSAHNVGTVDPGAGANNPDHRT 120

QY 121 YKIVRNNYKKONLHPYEDDYHNPRLHKFTVEAPIDMTSNMNGSTYSDRCKPBRVRIG 180
 DB 121 YKIVRNNYKKONLHPYEDDYHNPRLHKFTVEAPIDMTSNMNGSTYSDRCKPBRVRIG 180

QY 181 SGROFWRNDQDQVAVGAYHYLTAGTNHNRGANGSYLAGDVRKAGEYGLPIAGSK 240
 DB 181 SGROFWRNDQDQVAVGAYHYLTAGTNHNRGANGSYLAGDVRKAGEYGLPIAGSK 240

QY 241 GDSGSPMFTYDAEKOKMLNGILRSGNPEBKGNGFQJVRKSYFDEIFERDLTSLYTRA 300
 DB 241 GDSGSPMFTYDAEKOKMLNGILRSGNPEBKGNGFQJVRKSYFDEIFERDLTSLYTRA 300

QY 301 GNGVITISGNDGSGSITOKSGPSEIKITLANSPLKEXKYNPRYDGNYSPLN 360
 DB 301 GNGVITISGNDGSGSITOKSGPSEIKITLANSPLKEXKYNPRYDGNYSPLN 360

QY 361 NGEIYFMDOKQSLIFASDINOAGGLYFSGNFTVSPNSOTWOGAGIHVENSITVWK 420
 DB 361 NGEIYFMDOKQSLIFASDINOAGGLYFSGNFTVSPNSOTWOGAGIHVENSITVWK 420

QY 421 VNGVEHRLSKTGKTLHVOAKGKNGSISVGDGKYLLEQADQGNQASGLVYSGR 480
 DB 421 VNGVEHRLSKTGKTLHVOAKGKNGSISVGDGKYLLEQADQGNQASGLVYSGR 480

QY 481 GTVQVNDKQDPTDKEFVSGRGLDNLNHSITFKRIQNTDEGAMVNHNTQVANTTE 540
 DB 481 GTVQVNDKQDPTDKEFVSGRGLDNLNHSITFKRIQNTDEGAMVNHNTQVANTTE 540

QY 541 GNEIYVPLNGNNINKLTYRKEIAYNGMGEITDKNKNRNLNLYPTEDETLILSGGN 600
 DB 541 GNEIYVPLNGNNINKLTYRKEIAYNGMGEITDKNKNRNLNLYPTEDETLILSGGN 600

QY 601 LKEDITQTKGLFSGGRTPHAYNHLNKAHSMGEGITVDDHDMINRFRKAEVQIK 660
 DB 601 LKEDITQTKGLFSGGRTPHAYNHLNKAHSMGEGITVDDHDMINRFRKAEVQIK 660

QY 661 GSAVYRSVSSIEGNTMTVSNNANATFVGVPOQNTCTSPWGTGLTCOKVDELTKYI 720
 DB 661 GSAVYRSVSSIEGNTMTVSNNANATFVGVPOQNTCTSPWGTGLTCOKVDELTKYI 720

QY 721 NSIPKQINGSINTLTDNAITANYGLAKLKNVTLTNHSOFTLSNNATOIGNIRLSNSTA 780
 DB 721 NSIPKQINGSINTLTDNAITANYGLAKLKNVTLTNHSOFTLSNNATOIGNIRLSNSTA 780

QY 781 TVDNANLNGNVEHTLSAQFSLKNSHSSQIQDCKTIVLENAVTTHMSOFTLONLTAN 840
 DB 781 TVDNANLNGNVEHTLSAQFSLKNSHSSQIQDCKTIVLENAVTTHMSOFTLONLTAN 840

QY 841 STTINSAYSASANNTPR-PRSLSETPTPTSREHRENTLVNKGKSGOCTQFISLFG 899
 DB 841 STTINSAYSASANNTPR-PRSLSETPTPTSREHRENTLVNKGKSGOCTQFISLFG 899

QY 899 KSDTKLKNDAEDYILSVRNTGKEBETLEQITVESKDNQPLSDKLFTEINDHVDAG 959
 DB 899 KSDTKLKNDAEDYILSVRNTGKEBETLEQITVESKDNQPLSDKLFTEINDHVDAG 959

QY 959 LRYLYKNGCEERLHNPKEOELHNDLVABAEKTLHAQVFPKATQTEPERYRRA 1019
 DB 959 LRYLYKNGCEERLHNPKEOELHNDLVABAEKTLHAQVFPKATQTEPERYRRA 1019

QY 1019 WRYELVKNNGEERLHNPKEOELHNDLVABAEKTLHAQVFPKATQTEPERYRRA 1019
 DB 1019 WRYELVKNNGEERLHNPKEOELHNDLVABAEKTLHAQVFPKATQTEPERYRRA 1019

QY 1020 ARAAFPTLIPDQSLINLAKQAELEAFQSKA-----KTKVSKR---AV 1064
 DB 1020 ARAAFPTLIPDQSLINLAKQAELEAFQSKA-----KTKVSKR---AV 1064

DB 1011 -RAAFSTPDPDQSLINLAKQAELELTNNQOVNAGAVONKATALNKNRQKTTQKANTV 1069
 QY 1065 FSDPLLDQSL-----FA-----LEALEYIDAPQOSEKDLAQBKQKQKD 1108
 DB 1070 LATAVKEKTAQIDFANAKLAQMLTQOLEKALAV---AEQAEKERRAQBKQKQKD 1126

QY 1109 LLSRYSNSALSELATVNSMCLSVODELRLFVDQASAVNTNINQDERRDPAFAYQ 1168
 DB 1127 LLSRYSNSALSELATVNSMCLSVODELRLFVDQASAVNTNINQDERRDPAFAYQ 1185

QY 1169 OKTNLROGVQKALANGRIGAVFSHRSKDNTPEDQYKHAULTMAGSGPAQYQMGDLQFQV 1228
 DB 1186 QCTNLRQGVQKALANGRIGAVFSHRSKDNTPEDQYKHAULTMAGSGPAQYQMGDLQFQV 1245

QY 1229 NVGTGISASMAEGRKTHRRANVGNASVQFRLQGLQIPFYGNRYFIERENYQSE 1288
 DB 1246 NVGTGISASMAEGRKTHRRANVGNASVQFRLQGLQIPFYGNRYFIERENYQSE 1305

QY 1289 ERYKVTBSLAPNRYNAGRVDTFTPTDNTSVKPYFNNVTVDSNNAVOTTVNLTVLQOP 1348
 DB 1306 ERYKVTBSLAPNRYNAGRVDTFTPTDNTSVKPYFNNVTVDSNNAVOTTVNLTVLQOP 1365

QY 1349 PERYQKEVGIKAEILHFOISAFISKSQSGQLKQONVGVKLYRW 1394
 DB 1366 PERYQKEVGIKAEILHFOISAFISKSQSGQLKQONVGVKLYRW 1411

RESULT 10
 PCT-US03-05226-13
 ; Sequence 13, Application PC/TUS0305226
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RT/DC/DBR
 ; CURRENT APPLICATION NUMBER: PCT/US03/05226
 ; CURRENT FILING DATE: 2003-02-18
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 13
 ; LENGTH: 1436
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 PCT-US03-05226-13

Query Match 79.8%; Score 5807; DB 1; Length 1436;
 Best Local Similarity 79.3%; Pred. No. 0;
 Matches 1149; Conservative 78; Mismatches 154; Indels 68; Gaps 14;

QY 1 MKKTVRLNPLFLACISLGIIVSQAAGHTYFGIDYQYRDPANRKGKFTVGAONIKYVNRQ 60
 DB 1 MKKTVRLNPLFLACISLGIIVSQAAGHTYFGIDYQYRDPANRKGKFTVGAONIKYVNRQ 60

QY 61 GQVGTSMTKAPMIDSVSVSRNGVAALYENQYIVSAHNVGTVDPGAGANNPDHRT 120
 DB 61 GQVGTSMTKAPMIDSVSVSRNGVAALYENQYIVSAHNVGTVDPGAGANNPDHRT 120

QY 121 YKIVRNNYKKONLHPYEDDYHNPRLHKFTVEAPIDMTSNMNGSTYSDRCKPBRVRIG 180
 DB 121 YKIVRNNYKKONLHPYEDDYHNPRLHKFTVEAPIDMTSNMNGSTYSDRCKPBRVRIG 180

QY 181 SGROFWRNDQDQVAVGAYHYLTAGTNHNRGANGSYLAGDVRKAGEYGLPIAGSK 240
 DB 181 SGROFWRNDQDQVAVGAYHYLTAGTNHNRGANGSYLAGDVRKAGEYGLPIAGSK 240

QY 241 GDSGSPMFTYDAEKOKMLNGILRSGNPEBKGNGFQJVRKSYFDEIFERDLTSLYTRA 300
 DB 241 GDSGSPMFTYDAEKOKMLNGILRSGNPEBKGNGFQJVRKSYFDEIFERDLTSLYTRA 300

QY 295 SLTYRAGVYTTISGNDGQSSITQ---KSGIPSEIKITLANNSLPEKXKHNRV-- 349
 DB 301 TLEPRNSHSTPTNNNGCTVYQNEKVSMPQKRYQVQNEALKEKXK--EPVYAA 358
 QY 350 DGPVYSPRLNNGETLYFMDQKQSLIFASDINQAGGLYFEGNFTVSPNSNOTWQAGI 409
 DB 359 GGVYAYKPRLNNGKTIYFGDRGTGTLTENNINQAGGLYFEGNFTVSENNATWQAGV 418
 QY 410 HVSNSVTVMKVNGVEHDLRSKI GKGLTFVQAKGENKSSISVGDQVYI..EQOADDQGNQ 469
 DB 419 HVBDSVTVMKVNGVEHDLRSKI GKGLTFVQAKGENKSSISVGDQVYI..EQOADDQGNQ 478
 QY 470 AFSEIGLVSGGTQVQANDKQFDTDKFYFGSGGGLDNGHSTIFEGTINTDEBAMTVNH 529
 DB 479 AFKEGVIVSGRATVQJNSAQVDPNNIYFGFGRGLDNGHSTIFEGTINTDEBAMTVNH 538
 QY 530 NTTOANVTITGNES-VLP-NGNNINKLDYRKEIAYNGWFEGTIDKNGKNGRLNLTPEPT 588
 DB 539 NTTOANVTITGNES-VLP-NGNNINKLDYRKEIAYNGWFEGTIDKNGKNGRLNLTPEPT 598
 QY 589 EDRLLASGGTNLKGDTITQTKGLFSGSRPTPHAYNHLAKRWSMEGLFOGRIVWHDWI 648
 DB 599 EDRLLASGGTNLKGDTITQTKGLFSGSRPTPHAYNHLAKRWSMEGLFOGRIVWHDWI 656
 QY 649 NRTKANEQVIGSSAVSVSVSSIEGNTVSNNAKAFEGVTPNQNTICTRSQWGLTT 708
 DB 657 TETKANEQVIGSSAVSVSVSSIEGNTVSNNAKAFEGVTPNQNTICTRSQWGLTT 716
 QY 709 CQVYLDVDTKINS-EFKTQINGSLNTDNTANVKGAKLNGVNTLTINHSQFTLNNATQ 768
 DB 717 CKVYLDVDTKINS-EFKTQINGSLNTDNTANVKGAKLNGVNTLTINHSQFTLNNATQ 776
 QY 769 IGNRLSDNSTATVDNANLNGVNLTDQAQSLKXSHFSHQIQGDKGTTVLENATVMP 828
 DB 777 TGNQLSNHATVNNATVNGVNLTDQAQSLKXSHFSHQIQGDKGTTVLENATVMP 836
 QY 829 SDPTLQNTLNNSTLTNSAYSASNNTPR-RSGFTTPTTSAHRPNTLVNGLXISQ 887
 DB 837 SDATLQNTLNNSTLTNSAYSASNNTPR-RSGFTTPTTSAHRPNTLVNGLXISQ 896
 QY 888 STFOFTSS-FGVKSDKLLXNDAGDYILSVNTEKEPETLEQTLVESKONQPSDKL 947
 DB 897 STFOFTSS-FGVKSDKLLXNDAGDYILSVNTEKEPETLEQTLVESKONQPSDKL 956
 QY 948 FTLENDHVDAGALRYKLVNDGEFRLANPIKEQLHNDLVPAQAEETLEAKYBPFAKT 1007
 DB 957 FTLENDHVDAGALRYKLVNDGEFRLANPIKEQLHNDLVPAQAEETLEAKYBPFAKT 1016
 QY 1008 QTGPVKYRSRAARAAPBPTLPDQSLNMLEAKQELFAETQSKXAKTKQKRSRAVPSF 1067
 DB 1017 QTSNARVRSK--RAVFSDTLPDQSLQVQAEQVEPTAEKQKKA--EKVASKRAVPSF 1071
 QY 1068 PILDQS-----LFALEALEVID 1085
 DB 1072 TLEPQSQDLVQAEQVEPTAEKQKKAARARPSPTPLDLSKLTLEKLVIN 1131
 QY 1086 APQSEKORLAQEAERKQKDLISRYNSALSTLSATVNSMLSVQDELRLPYDQAS 1145
 DB 1133 AQQVQKKE---PDQEKORQKQDLISRYNSALSTLSATVNSMLSVQDELRLPYDQAS 1188
 QY 1146 AVMTNIAQDKRARDYDAFRAYQOQKNTLQIQVOKALANGIAGVSHSRNTPEDQV 1205
 DB 1189 AVMTNIAQDKRARDYDAFRAY-QQKNTLQIQVOKALANGIAGVSHSRNTPEDQV 1247
 QY 1206 NHATLTMSSGFAQYQMDLQFQVNVGTGISASMAEBSRKIRKAINGVNVAQYORLG 1265
 DB 1248 NHATLTMSSGFAQYQMDLQFQVNVGTGISASMAEBSRKIRKAINGVNVAQYORLG 1307
 QY 1266 QLGQIPYFGNRTFERENYQSEEVYKTPSLAEFRNNAIGRDVYPTTNDISIKYEPF 1325
 DB 1308 QLGQIPYFGNRTFERENYQSEEVYKTPSLAEFRNNAIGRDVYPTTNDISIKYEPF 1367
 QY 1326 VNYVDVSNANVQTTNLTVLQQPFGRVWKEVGLAKAELHFOJSAFISKQSQSLQKQON 1385

DB 1368 VNYVDVSNANVQTTNLTVLQQPFGRVWKEVGLAKAELHFOJSAFISKQSQSLQKQON 1427
 QY 1386 VGVXKGYRM 1394
 DB 1428 VGVXKGYRM 1436
 RESULT 11
 US-10-687-046-13
 ; Sequence 13, Application US/10687046
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-5941-1/RET/DCF/DER
 ; CURRENT APPLICATION NUMBER: US/10/687,046
 ; PRIOR FILING DATE: 2003-10-15
 ; PRIOR APPLICATION NUMBER: US/10/080,505
 ; PRIOR FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 13
 ; LENGTH: 1436
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 US-10-687-046-13
 Query Match 79.8%; Score 5807; DB 31; Length 1436;
 Best Local Similarity 79.3%; Pred. No. 0;
 Matches 1149; Conservative 78; Mismatches 154; Indels 68; Gaps 14;
 QY 1 MKKTVFLNPLFACISIGIVSOAMAGTYRGIDYQYRDPFAENKSKETVGAQNIKYVQ 60
 DB 1 MKKTVFLNPLFACISIGIVSOAMAGTYRGIDYQYRDPFAENKSKESVGAQNIKYVNE 60
 QY 61 GQVIGSTMTKAPMIDFSVSVSRNGVAALVENQYIVSAHNVGYTVDVDEAGSNPDDHRT 123
 DB 61 GQVIGSTMTKAPMIDFSVSVSRNGVAALVENQYIVSAHNVGYTVDVDEAGSNPDDHRT 120
 QY 121 YKIVRNRYKKNLHPEDDYRNPRHKEVTEAPIDMTSNNGSYSDSKYPRSVRIG 180
 DB 121 YQIVRNRYKKNKCNPHDHYMPRLHKEVTDAPAKMTNNNGKNTADSKYPRSVRIG 180
 QY 181 SGRQFPRNDQD-----KQDVAGAYHYLLTAGTTHNRGAGNGSYLAGDVYKAGBYGPP 235
 DB 181 TSGQWKTDEQKQKSSKSLADATVRLKNGTHSGAGNGVTVNLSGDLTKNNVGP 240
 QY 236 IAGSKDSGSPMTYDAEYQKMLINGILRBGNPBGKNGFQYVRSYF-DEIFERDLHT 294
 DB 241 TGVSFEDSGSPMTYDAIKQKMLINGVLTQNGNPFSGAGNGFOLIRKMWFDVNFVEDLEI 300
 QY 295 SLTYRAGVYTTISGNDGQSSITQ---KSGIPSEIKITLANNSLPEKXKHNRV-- 349
 DB 301 TLEPRNSHSTPTNNNGCTVYQNEKVSMPQKRYQVQNEALKEKXK--EPVYAA 358
 QY 350 DGPVYSPRLNNGETLYFMDQKQSLIFASDINQAGGLYFEGNFTVSPNSNOTWQAGI 409
 DB 359 GGVYAYKPRLNNGKTIYFGDRGTGTLTENNINQAGGLYFEGNFTVSENNATWQAGV 418
 QY 410 HVSNSVTVMKVNGVEHDLRSKI GKGLTFVQAKGENKSSISVGDQVYI..EQOADDQGNQ 469
 DB 419 HVBDSVTVMKVNGVEHDLRSKI GKGLTFVQAKGENKSSISVGDQVYI..EQOADDQGNQ 478
 QY 470 AFSEIGLVSGGTQVQANDKQFDTDKFYFGSGGGLDNGHSTIFEGTINTDEBAMTVNH 529
 DB 479 AFKEGVIVSGRATVQJNSAQVDPNNIYFGFGRGLDNGHSTIFEGTINTDEBAMTVNH 538
 QY 530 NTTOANVTITGNES-VLP-NGNNINKLDYRKEIAYNGWFEGTIDKNGKNGRLNLTPEPT 588
 DB 539 NTTOANVTITGNES-VLP-NGNNINKLDYRKEIAYNGWFEGTIDKNGKNGRLNLTPEPT 598

Db 539 NTQVANITITNESITAPSNKNNIKLDSKEIAVNGFGETDENKNGRLNIYKPTT 598
 Qy 569 EDETLISGGTNLKGDITQKGLFSSGPPPAHNLAKSEMEGIPQGEIYWDHWI 648
 Db 599 EDETLISGGTNLKGNITQEGTLFSGPPPAHNLN-REPNEI-GRPOGEVIDDWI 656
 Qy 649 NRTPEANFOIKGSAVAVSRVSSIEGNWTSNNNAKTFGVVNOONTICRSPTGLT 708
 Db 657 TRTFEENFOIKGSAVAVSRVSSIEGNWTSNNNAKTFGVVNOONTICRSPTGLT 716
 Qy 709 CQKVDITPTKVINSIPTQJNGSINLTDNATANYGLAKINGVTLNHSQFTLSNNATQ 768
 Db 717 CKTVLIDPTKVINSIPTQJNGSINLTDNATANYGLAKINGVTLNHSQFTLSNNATQ 776
 Qy 769 IGNIKLSDNSTATVDNANLNGVNLIDSAQSLKNSHFSHQIQGKGTTVTLNATWMP 828
 Db 777 TGNILDSNNAATVDNANLNGVNLIDSAQSLKNSHFSHQIQGKGTTVTLNATWMP 836
 Qy 829 SDPTLQNTLANS-ITLNSAVSASSNTPR-RSE-ETETPTSAEHRNTLTVNGLSQ 887
 Db 837 SDATLQNTLANS-ITLNSAVSASSNTPR-RSE-ETETPTSAEHRNTLTVNGLSQ 896
 Qy 888 GTFPTSSILFGYKSDIKLSDNABDYLLSVRNTGKPELTLEQLTVESKDNQPLSDKX 947
 Db 897 GTFPTSSILFGYKSDIKLSDNABDYLLSVRNTGKPELTLEQLTVESKDNQPLSDKX 956
 Qy 948 FTLNDHYDAGALRYKLVKNGDEPFLNPIEGEELNDYRAEQERTLEAKYETPACT 1007
 Db 957 FTLNDHYDAGALRYKLVKNGDEPFLNPIEGEELNDYRAEQERTLEAKYETPACT 1016
 Qy 1008 QTGEPRVSRRAAALFEDTLPOSLNALNAEQALTAETQSKAKTKVRSKRAVSD 1067
 Db 1017 QTSNARVSK---RAVSPDLPOSLDVLQALQVEPTLEKQAKKA---KAVSKRAVSD 1071
 Qy 1068 PLIDQS-----LFALAEVLD 1085
 Db 1072 TLPOSLDVLQALQVEPTLEKQAKTKVRSKRAVSDPDLRLKLVLEVLN 1131
 Qy 1086 APQSEKRLAQEAEKOKKODLSRYSNALSSELSATVMSMSVDELDRLPTQAS 1145
 Db 1132 AQQQVKE---PQSEKOKKODLSRYSNALSSELSATVMSMSVDELDRLPTQAS 1188
 Qy 1146 AVTTLAQDKRYSDAFRAYQOQKTNLRQYQKALANGRLGAVFSSHSNDTPEQYK 1205
 Db 1189 AVTTLAQDKRYSDAFRAY-QQKTNLRQYQKALANGRLGAVFSSHSNDTPEQYK 1247
 Qy 1206 NHATLTMWSGFAGYQWGDLOFQVNVGTGISASMAEBSGRKIHRAINYVANASYQFRLG 1263
 Db 1248 NHATLTMWSGFAGYQWGDLOFQVNVGTGISASMAEBSGRKIHRAINYVANASYQFRLG 1307
 Qy 1266 QLGIOFYGVNRYFIRENYSSEBYKXTSLAFNRYNAGRNTYTPTDNLSYKPYEF 1325
 Db 1308 QLGIOFYGVNRYFIRENYSSEBYKXTSLAFNRYNAGRNTYTPTDNLSYKPYEF 1367
 Qy 1326 VNYVDVSNANVQTTVALTVLQOPFGRYWQKEVGLKALIFHCISAFISKQSSQSGKQON 1385
 Db 1368 VNYVDVSNANVQTTVALTVLQOPFGRYWQKEVGLKALIFHCISAFISKQSSQSGKQON 1427
 Qy 1386 VGVKGYRW 1394
 Db 1428 VGVKGYRW 1436

RESULT 12
 PCT-US03-05226-15
 ; Sequence 15, Application PC/TUS0305226
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; CURRENT APPLICATION NUMBER: PCT/US03/05226
 ; CURRENT FILING DATE: 2003-02-18
 ; PRIORITY APPLICATION NUMBER: US 08/256,791

; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 15
 ; LENGTH: 1391
 ; TYPE: ERT
 ; ORGANISM: Haemophilus influenzae
 ; PCT-US03-05226-15

Query Match 72.8%; Score 5303.5; DB 1; Length 1391;
 Best Local Similarity 74.8%; Pred. No. 0;
 Matches 1053; Conservative 109; Mismatches 215; Indels 31; Gaps 15;

Qy 1 MKCTVPLNLEFLACISLIGVSOANAHTYGGIDYQYRDFAPENKGRPTGQNIKYKQ 60
 Db 1 MKCTVPLNLEFLACISLIGVSOANAHTYGGIDYQYRDFAPENKGRPTGQADIDYK 60
 Qy 61 GQLVGTSMTRAMPIDFVSVSRGVVALVENVQYVSAHVNGYTDVDPGAEKNPPQHRFT 120
 Db 61 GEMIGTWKGVPMWPLDSNVRGYSTLISRHQLISVAHVNGYDVDPGAEKNPPQHRFX 120
 Qy 121 YKIKRNNYKQNLHPEDYDHNRLKRYTAAALPDMTSMNGSTYSRRTKYPERRVIG 180
 Db 121 YKIKRNNYKQNLHPEDYDHNRLKRYTAAALPDMTSMNGSTYSRRTKYPERRVIG 178
 Qy 122 YKVKARINYSGRQ---YNDYQHRLEKVTETAPLEWVSFMDNRYKPNQYPLRVWG 178
 Db 122 YKVKARINYSGRQ---YNDYQHRLEKVTETAPLEWVSFMDNRYKPNQYPLRVWG 178
 Qy 123 SGROFMRNDQK---GDQVAGAYHTLTAGTNHQAGANGSYLGGDVRKAGEYGLPIAG 238
 Db 123 SGROFMRNDQK---GDQVAGAYHTLTAGTNHQAGANGSYLGGDVRKAGEYGLPIAG 237
 Qy 179 SGHGMWMDQNNKTIDILAVGCSWLGAGNTPEDEGAGNGTLEANGVQPNKYGGLFAG 237
 Db 179 SGHGMWMDQNNKTIDILAVGCSWLGAGNTPEDEGAGNGTLEANGVQPNKYGGLFAG 237
 Qy 238 SKDSSGSEWYTDARKQYVNLINGLRBNPFECKENGQVLYRSYDELPEDLHTSLYT 298
 Db 238 SKDSSGSEWYTDARKQYVNLINGLRBNPFECKENGQVLYRSYDELPEDLHTSLYT 297
 Qy 239 SFQDSSGSEWYTDARKQYVNLINGLRBNPFECKENGQVLYRSYDELPEDLHTSLYT 297
 Db 239 SFQDSSGSEWYTDARKQYVNLINGLRBNPFECKENGQVLYRSYDELPEDLHTSLYT 297
 Qy 299 RAGNGYVTLISGNDG---QGSITQSGIPSEITITLANSI-PLKEDXKHNRYNDP 352
 Db 299 RAGNGYVTLISGNDG---QGSITQSGIPSEITITLANSI-PLKEDXKHNRYNDP 352
 Qy 353 TNAETFRNIGSOSHNRVATITSTLPKKAQPERIVGLVNSQLHARDNRGDSFSTYGP 357
 Db 353 TNAETFRNIGSOSHNRVATITSTLPKKAQPERIVGLVNSQLHARDNRGDSFSTYGP 357
 Qy 358 NIVSPLNNGETLYEMDQKQSLIFASDINGAGGLYEGNFTYSRNSQ-TMOCAGIHY 411
 Db 358 NIVSPLNNGETLYEMDQKQSLIFASDINGAGGLYEGNFTYSRNSQ-TMOCAGIHY 411
 Qy 358 NIVSPLNNGETLYEMDQKQSLIFASDINGAGGLYEGNFTYSRNSQ-TMOCAGIHY 417
 Db 358 NIVSPLNNGETLYEMDQKQSLIFASDINGAGGLYEGNFTYSRNSQ-TMOCAGIHY 417
 Qy 412 SENSTYTWKNGVEHDLRSLKIGKTLHQAQKENGKISYSDQKVTLEQADQDQKQAF 471
 Db 412 SENSTYTWKNGVEHDLRSLKIGKTLHQAQKENGKISYSDQKVTLEQADQDQKQAF 471
 Qy 418 GBSSTYTWKNGVEHDLRSLKIGKTLHQAQKENGKISYSDQKVTLEQADQDQKQAF 477
 Db 418 GBSSTYTWKNGVEHDLRSLKIGKTLHQAQKENGKISYSDQKVTLEQADQDQKQAF 477
 Qy 472 SEIGLVSGRTVQALNDQKQFPTQFGRGRGLDNGHSLTEKRLQNTDDEGAMTVNHT 531
 Db 472 SEIGLVSGRTVQALNDQKQFPTQFGRGRGLDNGHSLTEKRLQNTDDEGAMTVNHT 531
 Qy 478 XEYGVISGRATVQALNSADQVDPNNITYFGRGRGLDNGHSLTEKRLQNTDDEGAMTVNHT 537
 Db 478 XEYGVISGRATVQALNSADQVDPNNITYFGRGRGLDNGHSLTEKRLQNTDDEGAMTVNHT 537
 Qy 532 TOAANVTITNESIVLPNGNNINKLDYRKETLAVNGMFGETDKKNGRNLNLYPTTEDR 591
 Db 532 TOAANVTITNESIVLPNGNNINKLDYRKETLAVNGMFGETDKKNGRNLNLYPTTEDR 591
 Qy 538 SQTANITITGNATI---NSDSKQLNKKDIAFGMFGEDQKAKTNGRLNLYPTTEDR 593
 Db 538 SQTANITITGNATI---NSDSKQLNKKDIAFGMFGEDQKAKTNGRLNLYPTTEDR 593
 Qy 592 TLLISGTTNLKQDITQKGLFSSGPPPAHNLAKSEMEGIPQGEIYWDHWI 651
 Db 592 TLLISGTTNLKQDITQKGLFSSGPPPAHNLAKSEMEGIPQGEIYWDHWI 651
 Qy 594 HLLISGTTNLKQDITQKGLFSSGPPPAHNLAKSEMEGIPQGEIYWDHWI 653
 Db 594 HLLISGTTNLKQDITQKGLFSSGPPPAHNLAKSEMEGIPQGEIYWDHWI 653
 Qy 652 FKAENFOIKGSAVAVSRVSSIEGNWTSNNNAKTFGVVNOONTICRSPTGLT 711
 Db 652 FKAENFOIKGSAVAVSRVSSIEGNWTSNNNAKTFGVVNOONTICRSPTGLT 711
 Qy 654 FKAENFOIKGSAVAVSRVSSIEGNWTSNNNAKTFGVVNOONTICRSPTGLT 713
 Db 654 FKAENFOIKGSAVAVSRVSSIEGNWTSNNNAKTFGVVNOONTICRSPTGLT 713
 Qy 712 VDLIDTKVINSIPTQJNGSINLTDNATANYGLAKINGVTLNHSQFTLSNNATQ 771
 Db 712 VDLIDTKVINSIPTQJNGSINLTDNATANYGLAKINGVTLNHSQFTLSNNATQ 771
 Qy 714 VDLIDTKVINSIPTQJNGSINLTDNATANYGLAKINGVTLNHSQFTLSNNATQ 773
 Db 714 VDLIDTKVINSIPTQJNGSINLTDNATANYGLAKINGVTLNHSQFTLSNNATQ 773
 Qy 772 TRLSDNSTATVDNANLNGVNLIDSAQSLKNSHFSHQIQGKGTTVTLNATWMP 831
 Db 772 TRLSDNSTATVDNANLNGVNLIDSAQSLKNSHFSHQIQGKGTTVTLNATWMP 831
 Qy 774 IKLSNNAATVDNANLNGVNLIDSAQSLKNSHFSHQIQGKGTTVTLNATWMP 833
 Db 774 IKLSNNAATVDNANLNGVNLIDSAQSLKNSHFSHQIQGKGTTVTLNATWMP 833
 Qy 832 TLQNTLANSITLNSAVSASSNTPR-RSE-ETETPTSAEHRNTLTVNGLSQ 889
 Db 832 TLQNTLANSITLNSAVSASSNTPR-RSE-ETETPTSAEHRNTLTVNGLSQ 889

Db 834 TLONLTLNNSVTYLNKSAVSALSNAPRRRRESELETTPTTSASHPNTLTNGKLSGGGT 893
 QY 890 FQFTSSLPFGKSDKILKLSNDAEGDYLLSVRNTGKEPPTLQTLTVESKDNQPLSDXLKFT 949
 Db 894 FQFTSSLPFGKSDKILKLSNDAEGDYLLSVRNTGKEPPTLQTLTVESKDNQPLSDXLKFT 953
 QY 950 LENDVDAGALRYKLYKNDGEFRLNHPDKESGLHNDLVRAEQERTLEAKQVEPTAKTOT 1039
 Db 954 LENDVDAGALRYKLYKNDGEFRLNHPDKESGLHNDLVRAEQERTLEAKQVEPTAKTOT 1013
 QY 1010 GEKVRSRARARAAFPDLPDQSLNLNLEAKQELTETOKSRKTKKVRKRAV--FSD 1067
 Db 1014 SKAVRER--RAVSDPLPAQSLNLNLEAKQ-LTTEIQTSKA--KVRSRARARAFSD 1067
 QY 1068 PLDDQSLPALRALAVYIDAPQSEKRLAQBAAEK--QRKQDLSRYSNALSLSATVN 1126
 Db 1068 TLDPQ--LLOALALEVIDAQQVKEPQTOBEERERQKCKE--SRYSNALSLSATVN 1124
 QY 1127 SMLSVODELDRLEFVDOASAVMTNIAODKRRYDSDAFRAYQQQKTLTQIGVOKALNGR 1186
 Db 1125 SMLSVODELDRLEFVDOASAVMTNIAODKRRYDSDAFRAY--QQKTLTQIGVOKALNGR 1183
 QY 1187 IGAVFSHSRSDNTFDEQVKNAHLLTNMGSFAQYQWGLQFGVAVGTGASAKMAEESGRK 1246
 Db 1184 IGAVFSHSRSDNTFDEQVKNAHLLTNMGSFAQYQWGLQFGVAVGTGASAKMAEESGRK 1243
 QY 1247 IHRKAIYGVNASYGFRLGQLGLOPFVGNRYZIERENYSEVRYKTPSLAENRYAGI 1306
 Db 1244 IHRKAIYGVNASYGFRLGQLGLOPFVGNRYZIERENYSEVRYKTPSLAENRYAGI 1303
 QY 1307 RVDYTFPTDNLISVYPFVYVYVDSNANVOTTVNLTVLOQPGRYWKEVGLAKETLHF 1366
 Db 1304 RVDYTFPTDNLISVYPFVYVYVDSNANVOTTVNLTVLOQPGRYWKEVGLAKETLHF 1363
 QY 1367 QISAFISKQSQSLGKQONVVKLGFRM 1394
 Db 1364 QISAFISKQSQSLGKQONVVKLGFRM 1391
 RESULT 13
 US-10-687-046-15
 : Sequence 15, Application US/10687046
 : GENERAL INFORMATION:
 : APPLICANT: St. Gene, Joseph W.
 : TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 : FILE REFERENCE: A-59941-1/RT/DOF/DHR
 : CURRENT APPLICATION NUMBER: US/10/687,046
 : PRIOR FILING DATE: 2003-10-15
 : PRIOR APPLICATION NUMBER: US/10/080,505
 : PRIOR FILING DATE: 2002-32-22
 : PRIOR APPLICATION NUMBER: US 08/296,791
 : PRIOR FILING DATE: 1994-10-25
 : PRIOR APPLICATION NUMBER: US 09/839,996
 : PRIOR FILING DATE: 2001-04-20
 : NUMBER OF SEQ ID NOS: 58
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 15
 : LENGTH: 1391
 : TYPE: PRT
 : ORGANISM: Haemophilus influenzae
 US-10-687-046-15
 Query Match 72.9%; Score 5303.5; DB 31; Length 1391;
 Best Local Similarity 74.8%; Pred. No. 0;
 Matches 1053; Conservative 109; Mismatches 215; Indels 31; Gaps 15;

QY 221 YKIVARNYKXKMLHPEEDYHNPELHKFYTEAPIDMTSNMNGSYSDRTKYPEKVRIG 180
 Db 221 YKIVARNYKXKMLHPEEDYHNPELHKFYTEAPIDMTSNMNGSYSDRTKYPEKVRIG 178
 QY 261 SGRQFMRDQOK--GQVAGAVHITLQNTNHRKQAGNGSYSLGQVRAKGEYGLPLNG 238
 Db 279 SGRQFMRDQOK--GQVAGAVHITLQNTNHRKQAGNGSYSLGQVRAKGEYGLPLNG 237
 QY 239 SKGDSGSPMPFYDAEQKWLINGLREGNPFEEKNGFOLYKSYFDEIPEERDLHTSLYT 298
 Db 238 SKGDSGSPMPFYDAEQKWLINGLREGNPFEEKNGFOLYKSYFDEIPEERDLHTSLYT 297
 QY 299 RAGNGVYITSGNDG----QGSITQSGIPSEIKITLANSI--PLKXKQVKNPRYDGP 352
 Db 298 RAGNGVYITSGNDG----QGSITQSGIPSEIKITLANSI--PLKXKQVKNPRYDGP 357
 QY 353 NITYSPRLNNGEITLPMDOQKSLIPASDINQAGGIFPFGNFTYSRPSNQ--TMOGAGHY 411
 Db 358 NITYSPRLNNGEITLPMDOQKSLIPASDINQAGGIFPFGNFTYSRPSNQ--TMOGAGHY 417
 QY 412 SENSFTVTKVGVYEHRLSKIGKGLHVAQKNGKSIYSGDKVLEQDADQKQKAF 471
 Db 413 SENSFTVTKVGVYEHRLSKIGKGLHVAQKNGKSIYSGDKVLEQDADQKQKAF 477
 QY 472 SEIGVSGRGTVQUNDKQFDTKFYGFGRLDNGSLTFKRLQNTDEGAMVNT 531
 Db 479 SEIGVSGRGTVQUNDKQFDTKFYGFGRLDNGSLTFKRLQNTDEGAMVNT 537
 QY 532 TOANVTITNGESIVLPMKNNINKLDYRKELAVNGFEGTQKNNKRLILYKPTEDR 591
 Db 538 TOANVTITNGESIVLPMKNNINKLDYRKELAVNGFEGTQKNNKRLILYKPTEDR 593
 QY 592 TLILSGTINKGDIOTKSKLFFSGRPTPHAYNHLKRWSEMEGIPQGLTWMDHWIKT 651
 Db 594 TLILSGTINKGDIOTKSKLFFSGRPTPHAYNHLKRWSEMEGIPQGLTWMDHWIKT 653
 QY 652 PRAENFOIKGGSANVSANVSIEGKMTVSNANATGVPVPOQNTICTRSDTGLITCOK 711
 Db 654 PRAENFOIKGGSANVSANVSIEGKMTVSNANATGVPVPOQNTICTRSDTGLITCOK 713
 QY 712 VDLTQTKYNSIPETQINGKINLTDNATVAVKGLATNGNVTLVNHSQFTLSNNAIOJEN 771
 Db 714 VDLTQTKYNSIPETQINGKINLTDNATVAVKGLATNGNVTLVNHSQFTLSNNAIOJEN 773
 QY 772 IRLSNSTAVVDANANNGVHLTDSAPSLKNSHFSHQIQGDKGTVLTENATWMPST 831
 Db 774 IRLSNSTAVVDANANNGVHLTDSAPSLKNSHFSHQIQGDKGTVLTENATWMPST 833
 QY 832 TLONLTLNNSVTYLNKSAVSALSNAPRRRRESELETTPTTSASHPNTLTNGKLSGGGT 889
 Db 834 TLONLTLNNSVTYLNKSAVSALSNAPRRRRESELETTPTTSASHPNTLTNGKLSGGGT 893
 QY 890 FQFTSSLPFGKSDKILKLSNDAEGDYLLSVRNTGKEPPTLQTLTVESKDNQPLSDXLKFT 949
 Db 894 FQFTSSLPFGKSDKILKLSNDAEGDYLLSVRNTGKEPPTLQTLTVESKDNQPLSDXLKFT 953
 QY 950 LENDVDAGALRYKLYKNDGEFRLNHPDKESGLHNDLVRAEQERTLEAKQVEPTAKTOT 1009
 Db 954 LENDVDAGALRYKLYKNDGEFRLNHPDKESGLHNDLVRAEQERTLEAKQVEPTAKTOT 1013
 QY 1010 GEKVRSRARARAAFPDLPDQSLNLNLEAKQELTETOKSRKTKKVRKRAV--FSD 1067
 Db 1014 SKAVRER--RAVSDPLPAQSLNLNLEAKQ-LTTEIQTSKA--KVRSRARARAFSD 1067
 QY 1068 PLDDQSLPALRALAVYIDAPQSEKRLAQBAAEK--QRKQDLSRYSNALSLSATVN 1126
 Db 1068 TLDPQ--LLOALALEVIDAQQVKEPQTOBEERERQKCKE--SRYSNALSLSATVN 1124
 QY 1127 SMLSVODELDRLEFVDOASAVMTNIAODKRRYDSDAFRAYQQQKTLTQIGVOKALNGR 1186
 Db 1125 SMLSVODELDRLEFVDOASAVMTNIAODKRRYDSDAFRAY--QQKTLTQIGVOKALNGR 1183

QY 1187 IGAVFHSRSDNTFDEOVKNHATLTMWGSFAYOMGDLQFQVNVNGNISAKVABOSRK 1246
 DB 1184 IGAVFHSRSDNTFDEOVKNHATLTMWGSFAYOMGDLQFQVNVNGNISAKVABOSRK 1243
 QY 1247 IHRKAINGVNASYQFRLGOLGIPYGVNRYFERENYQSEEVVKTPLAFRNVNAGI 1306
 DB 1244 IHRKAINGVNASYQFRLGOLGIPYGVNRYFERENYQSEEVVKTPLAFRNVNAGI 1303
 QY 1307 RVYFTPTDNIISVKEPFYVNVYDVSNANVQTVNLVLOQFGRYMKVEVGLKAEILHF 1366
 DB 1304 RVYFTPTDNIISVKEPFYVNVYDVSNANVQTVNLVLOQFGRYMKVEVGLKAEILHF 1363
 QY 1367 QISAFISKSQSGOLGKQOONVGVKLYGM 1394
 DB 1364 QISAFISKSQSGOLGKQOONVGVKLYGM 1391

RESULT 14

PCT-US03-05226-11
 Sequence 11, Application PC/TUS0305226
 GENERAL INFORMATION:
 APPLICANT: St. Geme, Joseph W.
 TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 FILE REFERENCE: A-59941-1/RFT/DCF/DHR
 CURRENT APPLICATION NUMBER: PCT/US03/05226
 PRIOR FILING DATE: 2003-02-28
 PRIOR FILING DATE: 1994-10-25
 PRIOR APPLICATION NUMBER: US 09/839,996
 PRIOR FILING DATE: 2001-04-20
 NUMBER OF SEQ ID NOS: 58
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 11
 LENGTH: 1391
 TYPE: PRT
 ORGANISM: Haemophilus influenzae
 PCT-US03-05226-11

Query Match 72.8%; Score 5297.5; DB 1; Length 1391;
 Best Local Similarity 74.7%; Pred. No. 0;
 Matches 1052; Conservative 109; Mismatches 216; Indels 31; Gaps 15;

QY 1 MKKTVRLNFLTACISGIVSOMAGHTYFGIDYQYRDPFAENKCKFTVGNOMIKYNNQ 60
 DB 1 MKKTVRLNFLTACISGIVSOMAGHTYFGIDYQYRDPFAENKCKFTVGAQDIDYNNK 60
 QY 61 GQIVGTSMTKAPMIDFVSVRNGVALVENQYIVSAHVNVGYTDVDFGABENPDQHRFT 120
 DB 61 GEMIGTMKGVPMEDLSVRVGGYSTLLSBQHLSTVAHVNVGYDVIDFGBENPDQHRFX 120
 QY 121 YKTVRNNYKKNLHPREDYHNPRIKFTYEAAPIDMTSNMNGSTYSDRKTPREVRIG 180
 DB 121 YKTVRNNYKKNLHPREDYHNPRIKFTYEAAPIDMTSNMNGSTYSDRKTPREVRIG 178
 QY 181 SGRQFENDDOOK--GDQVAGAYHYLTAGNTNNGAGANGSVLGGDVAKAGEPLPIHG 238
 DB 179 SGRQMKCDANKKTIGDLYAG--SWLIGNTFEDGPRANGTLELNGRVNENKKGPELPIHG 237
 QY 239 SKDGSQPMFTYDAEKOKMLINGILRBGNPEGEKNGAFOLVKYSYDFEIERDLATSLYT 298
 DB 238 SFGDSSPMFTYDAEKOKMLINGILRBGNPEGEKNGAFOLVKYSYDFEIERDLATSLYT 297
 QY 299 RAGNGYTTISGUNG-----QSSITQSGIPESEKITTANMGL-PLKCKDKVNNPXYDOP 352
 DB 298 TNAEYFNGISDHNKRAVATIKSTLPERKAIOPERIVSLJDNQSILHARDKNDSPEYKXP 357
 QY 353 NISYPLNKGELTYNDOKOSLITASDINGAGAGLYFGNNTVSPNSNO--TWQAGAHY 411
 DB 358 NISYPLNKGELTYNDOKOSLITASDINGAGAGLYFGNNTVSPNSNO--TWQAGAHY 417
 QY 412 SENSTYTKVNGVEHDLKICGKGLTJVQAKGENKSGISVGGKYLIEQADDOGNKQAF 471
 DB 418 GBSSTYTKVNGVEHDLKICGKGLTJVQAKGENKSGISVGGKYLIEQADDOGNKQAF 477

QY 472 SETIGISRGTVLNDKQFDTOKFYFGFGRGLDINGHSLTPRRIONTDEGAMIVNNT 531
 DB 478 KKVIGVSGRATVQINADQVDPNNIYFGFGRGLDINGHSLTPRRIONTDEGAMIVNNA 537
 QY 532 TQAAVITTNESIVLPHGNINIKLDRKELANGPEPDKNGHNRILLYKPTTDR 591
 DB 538 TQAAVITTNESIVLPHGNINIKLDRKELANGPEPDKNGHNRILLYKPTTDR 593
 QY 592 TLLSGGTNKGDLITQNGKLFPSGRPTPHAVNHLNKEWSEMEGIPQELIWMHDMINT 651
 DB 594 HLLSGGTNKGDLITQNGKLFPSGRPTPHAVNHLNKEWSEMEGIPQELIWMHDMINT 653
 QY 652 FPAENFOLKGSAYVSRVSSIBGNWTVSNANATFGVPPQOANTICTRSDWTGLTTCQK 711
 DB 654 FPAENFOLKGSAYVSRVSSIBGNWTVSNANATFGVPPQOANTICTRSDWTGLTTCQK 713
 QY 712 VDLDTKVINSLPKQINGSLNLTDNA TAVKGLATNANTLTNHSOFTLSNNAVQGN 771
 DB 714 VDLDTKVINSLPKQINGSLNLTDNA TAVKGLATNANTLTNHSOFTLSNNAVQGN 773
 QY 772 IRLSDNSTATVFNANINGVHLTDSAFSLKNSHFHQIQGDKGTVTLLENATWTMPSDT 831
 DB 774 IRLSDNSTATVFNANINGVHLTDSAFSLKNSHFHQIQGDKGTVTLLENATWTMPSDT 833
 QY 832 TLQNTLTNASTTLNATSAISSNTP--RRSLFTTITTSAAHRRFVTLTVNGKLSGGGT 889
 DB 834 TLQNTLTNASTTLNATSAISSNTP--RRSLFTTITTSAAHRRFVTLTVNGKLSGGGT 893
 QY 890 FQFTSSLFGKSKDKLNSDAGDYTLISVNNYKPELTLBOLTVESKKNOPISDKMFT 949
 DB 894 FQFTSSLFGKSKDKLNSDAGDYTLISVNNYKPELTLBOLTVESKKNOPISDKMFT 953
 QY 950 LENDVDAGALRYKLVKNDGSEFLNFIKEOELHNDLVRABQAEKLEAKOYEPTAKQT 1009
 DB 954 LENDVDAGALRYKLVKNDGSEFLNFIKEOELHNDLVRABQAEKLEAKOYEPTAKQT 1013
 QY 1010 GEPKYSRRAPARAPEPTLIPQSLINALEKQALTAITOKSAKTKKYSKRAY--FSD 1067
 DB 1014 GEPKYSRRAPARAPEPTLIPQSLINALEKQALTAITOKSAKTKKYSKRAY--FSD 1067
 QY 1068 PLTQOSLFALEALAEVTPARQOSEKDLAQEAEK--OKRODLISVNSALSELATVN 1126
 DB 1066 PLTQOSLFALEALAEVTPARQOSEKDLAQEAEK--OKRODLISVNSALSELATVN 1124
 QY 1127 SMLSGVDELDRLFVDDAQSAAVTNIAQDKARYDSDAFRAVYQOQKTNLROIGYOKALANG 1186
 DB 1125 SMLSGVDELDRLFVDDAQSAAVTNIAQDKARYDSDAFRAVYQOQKTNLROIGYOKALANG 1183
 QY 1187 IGAVFHSRSDNTFDEOVKNHATLTMWGSFAYOMGDLQFQVNVNGNISAKVABOSRK 1246
 DB 1184 IGAVFHSRSDNTFDEOVKNHATLTMWGSFAYOMGDLQFQVNVNGNISAKVABOSRK 1243
 QY 1247 IHRKAINGVNASYQFRLGOLGIPYGVNRYFERENYQSEEVVKTPLAFRNVNAGI 1306
 DB 1244 IHRKAINGVNASYQFRLGOLGIPYGVNRYFERENYQSEEVVKTPLAFRNVNAGI 1303
 QY 1307 RVYFTPTDNIISVKEPFYVNVYDVSNANVQTVNLVLOQFGRYMKVEVGLKAEILHF 1366
 DB 1304 RVYFTPTDNIISVKEPFYVNVYDVSNANVQTVNLVLOQFGRYMKVEVGLKAEILHF 1363
 QY 1367 QISAFISKSQSGOLGKQOONVGVKLYGM 1394
 DB 1364 QISAFISKSQSGOLGKQOONVGVKLYGM 1391

RESULT 15

US-10-687-046-11
 Sequence 11, Application US/10687046
 GENERAL INFORMATION:
 APPLICANT: St. Geme, Joseph W.
 TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 FILE REFERENCE: A-59941-1/RFT/DCF/DHR

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw mode.

Run on: June 23, 2004, 17:30:42 ; Search time 21 seconds

[without alignments]
3036.187 Million cell updates/sec

Title: US-10-645-655-2
Perfect score: 7273

Sequence: 1 MKKTVFRLNFTLACTISLGV.....SGSQSLGQGVKGLGK 1394

Scoring table:

Gapop 10.0, Gapext 0.5

Searched: 302849 seqs, 45738825 residues

Total number of hits satisfying chosen parameters: 302849

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending patents AA New:
1: /cgn2_6/prodata/1/paa/PCT_NEM_COMB.pep.*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1609	22.1	1694	1 PCT-US04-06615-4	Sequence 4, Appli
2	1529	21.3	1794	1 PCT-US04-07001-763	Sequence 763, App
3	295.5	4.1	1542	1 PCT-US04-07001-670	Sequence 670, App
4	285.5	3.5	1004	5 US-09-806-709-57	Sequence 57, Appl
5	285.5	3.5	1010	5 US-09-806-709-55	Sequence 55, Appl
6	283.5	3.9	1222	5 US-09-806-709-37	Sequence 37, Appl
7	283.5	3.9	1228	5 US-09-806-709-34	Sequence 34, Appl
8	283.5	3.9	1536	5 US-09-806-709-67	Sequence 67, Appl
9	280	3.8	1477	5 US-09-806-709-71	Sequence 71, Appl
10	277	3.8	977	5 US-09-806-709-53	Sequence 53, Appl
11	277	3.8	983	5 US-09-806-709-51	Sequence 51, Appl
12	275	3.8	1381	1 PCT-US04-07001-693	Sequence 693, App
13	275	3.8	1492	1 PCT-US04-07001-672	Sequence 672, App
14	269	3.7	1005	5 US-09-806-709-41	Sequence 41, Appl
15	269	3.7	1011	5 US-09-806-709-39	Sequence 39, Appl
16	246.5	3.4	969	5 US-09-806-709-32	Sequence 32, Appl
17	246.5	3.4	975	5 US-09-806-709-30	Sequence 30, Appl
18	245.5	3.4	1180	5 US-09-806-709-65	Sequence 65, Appl
19	245.5	3.4	1188	5 US-09-806-709-63	Sequence 63, Appl
20	245.5	3.3	1073	5 US-09-806-709-49	Sequence 49, Appl
21	245.5	3.3	1079	5 US-09-806-709-47	Sequence 47, Appl
22	245.5	3.3	1290	6 US-10-786-892-138	Sequence 138, App
23	239	3.3	915	5 US-09-806-709-35	Sequence 35, Appl
24	238	3.3	10203	6 US-10-724-972A-4098	Sequence 4098, Ap
25	237	3.3	1035	5 US-09-806-709-69	Sequence 69, Appl
26	235	3.2	1531	1 PCT-US03-32645-44	Sequence 44, Appl

27	232	3.2	1336	5 US-09-806-709-73	Sequence 73, Appl
28	229	3.1	1395	5 US-09-806-709-45	Sequence 45, Appl
29	229	3.1	1103	5 US-09-806-709-43	Sequence 43, Appl
30	228	3.1	1220	5 US-09-806-709-28	Sequence 28, Appl
31	228	3.1	1226	5 US-09-806-709-26	Sequence 26, Appl
32	221	3.0	306	1 PCT-US04-12717-56	Sequence 56, Appl
33	221	3.0	906	6 US-10-831-070-56	Sequence 56, Appl
34	218.5	3.0	915	1 PCT-US03-31057-3	Sequence 3, Appl
35	218	3.0	992	5 US-09-806-709-61	Sequence 61, Appl
36	218	3.0	998	5 US-09-806-709-59	Sequence 59, Appl
37	212	2.9	3636	6 US-10-724-972A-5942	Sequence 5942, Ap
38	205.5	2.8	3829	1 PCT-US03-38966-10	Sequence 10, Appl
39	204.5	2.8	574	1 PCT-US03-31057-73	Sequence 73, Appl
40	204.5	2.8	578	1 PCT-US03-31057-141	Sequence 141, App
41	204.5	2.8	579	1 PCT-US03-31057-140	Sequence 140, App
42	204.5	2.8	580	1 PCT-US03-31057-139	Sequence 139, App
43	204.5	2.8	581	1 PCT-US03-31057-138	Sequence 138, App
44	204.5	2.8	582	1 PCT-US03-31057-137	Sequence 137, App
45	204.5	2.8	583	1 PCT-US03-31057-136	Sequence 136, App

ALIGNMENTS

RESULT 1
PCT-US04-06615-4
Sequence 4, Application PC/TUS0406615
GENERAL INFORMATION:
APPLICANT: New England Medical Center Hospitals, Inc.
APPLICANT: Plant, Andrew G
TITLE OF INVENTION: Treatment of IgM Deposition Diseases
FILE REFERENCE: 28154/2068
CURRENT APPLICATION NUMBER: PCT/US04/06615
CURRENT FILING DATE: 2004-03-09
PRIOR APPLICATION NUMBER: US 60/453055
PRIOR FILING DATE: 2003-03-07
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patent version 3.2
SEQ ID NO 4
LENGTH: 1694
TYPE: PRT
ORGANISM: Haemophilus influenzae
PCT-US04-06615-4

Query Match	22.1%	Score 1609	DB 1	Length 1694
Best Local Similarity	26.5%	Pred. No. 1.5e-78		
Matches 479	Conservative 251	Mismatches 536	Indels 542	Gaps 52
QY	6	PLNLTPLCTSLGIVSQMAGHTTFCGIDYQYRDPENKGFYVGAONITYNKKQGLVG	65	
DB	6	PLNLTPLCTSLGIVSQMAGHTTFCGIDYQYRDPENKGFYVGAONITYNKKQGLVG	65	
QY	66	TSMTKA-PWIDFVSU-ENGVALVENQYVSVAH-NGYTDVFG-AGEG-----NPDQ	117	
DB	66	NVLPMGSLPMDFVSUVDVDRJATLILNPDYVGVKHSNGSELFGULNNMNNKNAHSH	125	
QY	118	RFT-----YKIVKRNXY-----KKONLPEYEDYRNLHKEVTEAPID-MTS	160	
DB	126	RIVSSEERKPYSEKNEYETKLNKAYATEDQKRRDYVWPLDVEVAPBASTA	185	
QY	161	NNNGSTYSIDTKPEPVRISSGRQFMNDDKDCQ-----VAGAHYTL	203	
DB	186	SSDAGTYNDNKKPAPVRLGSSGQF-----LYKKDNYSILNNHEVGNLKLGAAYTG	242	
QY	204	TGNTHNORAGAGGYSLC-----GVRAGEYGPLIAGSKDSSGSPFIYDAEKQW	257	
DB	243	INGTYKKNHNNGLIGFNSKEHSPKXILSQDPLTNVAVLDSSPLFIYDAEKQW	302	
QY	258	LINGTLRGNPFEGKE-NGFQIVKSYFDEIFRDLHTSLYTRANGVTTISGNDGGS	316	
DB	303	LFLGSYFPMAGYNNKSKQWQBNVITYRPEAKTVLDKDTAGSL-----TGS	345	

QY 317 IIOKSGIPSEIKITLANNMLPLAKKOKVHNPYD--GPIVSPRLANGETLYMNOQOSL 375
 DB 346 NIOYNNMPCCTKSVSNCS-----BSLVNDIPJSSQUTDSKNNKSGVTL--KSGSL 397
 QY 376 IFASDINOAGAGLYEGNFTVSPNSNQ-TMOAGIHYSENSITVTKANGVEHEDLSCKG 434
 DB 398 TLANNIDOGAGGLFPGDYEYKGTSDSTTKAGASVADGCTTMYNHPKSDHAKLGR 457
 QY 435 GTLLAVOKKGNKGS-SVGDGVTLIEQADDOCKAFSEGLVSGGVQVANDKOPDND 494
 DB 458 GTLLVEKGENKSLKVDGTVILKQADANNKVAFEQVIGVSGSTVANDKQVDEN 517
 QY 495 KFYGFGRGLDANGSLTFKRIONTDEGAMVNNITQOANTINGNESIYLR----- 549
 DB 518 SIYGFGRGLDANGSNLTFEHRITNDGRLVNNHNSKSTVITIGSLLTDENTITPY 577
 QY 550 ----- 549
 DB 578 NIDAPEDNPAFRITKDGQLYNLNENTYVALRKAGASTRSELPRNGSGSENNMLYMK 637
 QY 550 -----GNNIKLDYRKELIANGFPEBTOKKNGRLNLYKPTEDRTLLSGGTLNG 603
 DB 636 TSDAKANVNNHINERNNGNGFYGR-EEGKNGNGLNVTFKGSEQRNFLLTGTVLNG 696
 QY 664 DITQKAKLFFSGRPTPAVY--HLKKESEMEG-FOGELVMDHMINRTKKNPQING 661
 DB 667 DLKREKSTLFGSRPTPAADIAISTYKQDHFENNEVWEDMINKNKANINIVN 756
 QY 662 GSAYVS-RNVSSIEGNMTVSNNAATFGVVPNOQNTICTSMTGLTTCQKVDLDTKYI 720
 DB 757 MTLVSGRNVANITSNIYASDNKVIHY--KAGDVCVMSDVTGVITCTDKSLD-KAL 813
 QY 721 NSLPTQINGSLNLDNATVNGJAKLKNVTLTNSQTLTSSNNAQTGINLSDNSTA 780
 DB 814 KSFNATVSGVNLGSAHF-VLGKANLFGISGTGNSQ----- 851
 QY 751 TVNNALANGVHLTSAQFSLKNSHSHQIQDGGTVTLNNAVTMPSDTLQNTLUN 840
 DB 852 -----VMT-----MNSHMLTGDSDNVQNLK 875
 QY 841 SITLMSAYSASNNTPRRRSLETETPTSAHRENTLTVNGKSGOCTFOPTSSLFYKX 900
 DB 876 GHILNAQNDARKVT-----YNLTLYN-SLSGNSFYLLDLSNKO 916
 QY 901 SDKLKSNABGDYILSVRYTGEPEPILSQTLVSGDQPLSDKLTLENDHDAGL 960
 DB 917 GKVAVTKSATGNFTLVADKTGEF-TKELTLTFLD--SNATRNINVSIVGNTVLDGAW 973
 QY 961 RYLVKNDGEFRLNPIKEOELNELY----- 987
 DB 974 KYKLAVNGRYOLYNF--EVEKNGQVDTMTITPENNIQADVPSPNNBELAVETVYP 1031
 QY 988 -----RAEQAEKTE-----AKOVEFLAK--TQGE-- 1011
 DB 1032 PPAPATSEHETVAENSKESTVEKNEQDATETTAONGEVAEFAKSVYKANTQNEVA 1051
 QY 1012 -----PKYRSBARAF-----DTL 1028
 DB 1092 QSGSETEETOTTEIKETAKYKGEBAKVEKOEIOEAPVASTSPKOKAPREVSTDK 1151
 QY 1029 PDGSLNML-----EAKOELTATOKSA----- 1053
 DB 1152 VEBTQVQAPQOSTIVAAMEATSPNSKPALETOPPSKKNAPVPIVVSQKQENTTDP 1211
 QY 1054 -----KTKYRSKR-----AVSPDLDO----- 1072
 DB 1212 TEKEKAVETEKEIOEPPOVASQSPKQOSETVOPQAVLESENVPTVNAHEVQACQOT 1271
 QY 1073 -----SLFALEALVETAPQO----- 1089
 DB 1272 QISAIVSTKOPAPENSLINGSATALTETAEKSKPOTETMASTEDASCHANTVADNSVA 1331
 QY 1090 -----SEKORLAQEE-----AEKOROK----- 1107

DB 1332 NNSSESDKSRBRKRSISQPOETSAEHTTAASDTETTLAONKSKSPNRRSRVRSEPTV 1391
 QY 1108 -----DLISYNSALSLSA-----TVNSMLSVODELDELFLV--DOAQAVWT 1149
 DB 1392 TNGSRSRSTVALRDLTSTNNAVTSIDMAKAQFVALVGAQVSQHSOLEMNNNGQVNVV 1451
 QY 1150 NIAODRRYSDAPRAYCOOKTNLRQIGVOKALANG-RIGAFVSHSRSDNTPEQVQNSNA 1208
 DB 1452 SMTSMNENYSSQRRFSSKSQ-TQIGMQTISNNVQCGVFTYVNSNNPD-KASSKN 1509
 QY 1209 TLTMSGPAQY-----QMGDLQGVNVGTGISASMAEQRIRKALINVGNASYOPEL 1264
 DB 1510 TLAQVNFYSKYADNEM--YLGIDIGYKFGQSNLKTNNNAFAHHTAQFGLTRAKAFNL 1566
 QY 1265 GOLGQZFYGVNRYPIERENYQSEBEVRYKTPSLAFPRVNAGRVDYFTETDMSYKPYE 1324
 DB 1567 GNFETTPYGVRSYSSNANFALAKDRIKONF-SVYTAAPQVDSYTY-HLSEFVTPIL 1625
 QY 1325 FVNVYVSNNAVQTTNLTVLQOPFGRYMOKEVGLKAEILHPOISAFISKSGSQGLKQO 1364
 DB 1626 SARY-DTNGSGKLTNVQYDPAVNVENQOQYNNAGLTKYHVKLSLIGLTRYAKQAEKQK 1684
 QY 1385 NGVYLCGY 1392
 DB 1685 TALKLSE 1692

RESULT 2
 PCT-US04-07001-763
 ; Sequence 763, Application PC/TUS0407001
 ; GENERAL INFORMATION:
 ; APPLICANT: BKALETZ et al.
 ; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONYPEABLE H. INFLUENZAE
 ; FILE REFERENCE: 28335/36815A
 ; CURRENT APPLICATION NUMBER: PCT/US04/07001
 ; CURRENT FILING DATE: 2004-03-05
 ; PRIOR APPLICATION NUMBER: US 66/453,134
 ; PRIOR FILING DATE: 2003-03-06
 ; NUMBER OF SEQ ID NOS: 771
 ; SOFTWARE: Patceptin version 3.2
 ; SEQ ID NO 763
 ; LENGTH: 1794
 ; TYPE: ERT
 ; ORGANISM: H. influenzae
 PCT-US04-07001-763

Query Match 21.0% Score 1529; DB 1; Length 1794;
 Best Local Similarity 25.4%; Pred. No. 3,3e-74;
 Matches 466; Conservative 238; Mismatches 536; Indels 654; Gaps 60;

QY 6 FRLNPLTACISLIGIVSOAMAGHTYFGIDYVYRDEFAEKKGKFTVGAQNTIKVNNKQGLWG 65
 DB 6 FKLNFIALVAVALTLPYTEALVANDVDYQIFRDLAKKKGKFGVGAINTVEADNNNNLIG 65
 QY 66 TSATK-APMLDSVVS-RNGVIAALVENQYTVSAH-ANGTIDVDFG-AGEG-----NPD 115
 DB 66 SALPKDIPMIDSAVDKRLATLVPYVGVKGVNGVGEJHFNGLNKNMPPFGNSI 125
 QY 116 QHRFT-----YKIVRNNY-----KKDLH-----PYE----- 138
 DB 126 QHRVSWEBKRYTYEKNSSSELMKGTQMEKDKQYTSNKKVDPSELVGLVYEQNQ 185
 QY 139 ---DYNNPLKHFVTEAPID--MISNMGSTYSDRKTYPERVYIGSGHOF----- 185
 DB 186 KREDDYMPRLDKFTVAVAPTEASTISSDAGTYNDQNKYPAFRLDSGSGOFTYKKGSHYE 245
 QY 186 -----WRNDQKGCYVAGAYHTLTAGTHNORAGANGSYGL-----GD 224
 DB 246 LILFEKNEKKEIIRHMDVGDGINKLVGNAYTTLGAGTPYKVNHTDGLGFGDSDTEHDND 305
 QY 225 VRKAGEYGPLTAGSKGDSGSPMEITYDAEKOKMLINGILBEGNPFGEKE-NGFOLVRKSY 283

Db 306 PKELLRKPRLTAVAVLGDGSGPLFVYDGSKEKMLFLGAYDFWGYGKKXWQENNYKPOF 365
 QY 284 FDEIFERDLHTSLYTRANGVYITISAKDNGGSIITQSGSPISITKITLANNSLFLKXOK 343
 Db 366 AENILKCC--SAGLKGNTQY--NWTSGKNTSLISGSESLVDYD----- 408
 QY 344 VHNPRYDGNISYSPILNNGERTJYFMOKQGSITFASDINGAGGAYFEGNTYSPNS-NO 402
 Db 409 -----MKNLHGNVTF--ESSGNLTLNNMIDQAGGLFFEGDYEVAGTSEMT 454
 QY 403 TWGAGIIVSENSYVYKRVNGVDEHLSKIGKGLIAYQAKGNSKISVGDGKYLLEQDA 462
 Db 455 TWKAGISVAGKTYKWKVFNPOFRLAKTIGKGLIYEGGDKNSKLVAGDGYVIAKQOT 514
 QY 463 DDOGKQAFSEGLVSGRGVQVANDXOFDTDKYFBEFGGGLJLNGSLTFRKIONTDE 522
 Db 515 TT--GGHAFASVAGISGRSTYVANDNDVDRNSYFQFGRGLDANGNMLTFEHNID 572
 QY 523 GAMIVHNTTOAANVTITG--NRSIV-----LPNGNLI----- 553
 Db 573 GARLYVHNTNINNSITITGAGLITPQVITITFALTADDDNYVYPSIPRGKOLYFSMT 632
 QY 554 -----MKLDYRKEL-----AYNGMFG 569
 Db 633 CYKXYALXQSGSPLAEMPYSEKSDANWEMECQNDQAKYAVYINNRBMGFMGYFG 692
 QY 570 -ETDKKHNKGLNLIYETTEDRKLILSGGNLKGDILOTYKGLFSGRPPIPAVHLM- 627
 Db 693 EEAKTADQNGKLVNTSGSKDQRFLLTGNTNLNGELKYEKTLFISGRPIPHARDIANI 752
 QY 628 -----KMSMEGIPQGEITWDMHINRTFKAEFQIKGSAVVS--RNVSIEGWTS 680
 Db 753 SSTEKDKHFAE-----KNEVYEDDMIKRTIKANINVTNNAITLVSGNVESTITS 807
 QY 681 XNANATPGVVPNQNTLCTRSPTMGLITQKVDLITDKYINSIPQVQNGSINLTDNATA 740
 Db 808 NKAKYHIGY--KAGDYCVASDYTGATCAN-DLSTKALNSFNPTNLRGNVLTESA-- 862
 QY 741 NKYGLAKLNGVTLNHSOPTLSNNAIQIGNIRLSDNSTRTATDANAKNGVHLTSDAGFS 800
 Db 863 -----NFTLG-----KANLEGINSTENSGYN 884
 QY 801 LKNSHFSQIQDGKGTVLENAATWMPSDTLTONTLNSTITLINSAYSASSNTPRRR 860
 Db 885 LK-----ENSHVTLGNSDVQICLJLNGHILHNNSDPTK----- 919
 QY 861 SLETTPTLSAHRPNTLTNGKLSGOGTFOFTSSLFGYKSDKLTLSDAEGDYIISVYN 920
 Db 920 -----ETKYHTLNIS--NLGSGSFYVWDFTKQSGKVVYKSAKGFITLQVAN 967
 QY 921 TGEKEPTLEQLIVESKQNPJSDKAKFTLENCHVDGALKYVNDGFRJANFIKQ 980
 Db 968 KXGEBPNH-NELTLEFDS-SNATERSGANVSLANGVDRGAMSYTLKENSGRYVJHNEVER 1025
 QY 981 -----ELHNDYV----- 988
 Db 1026 RQVYDTPSIATANNQADVPYSNNHEFTARVADPIPLAPAPAGSMAHEOPETRP 1085
 QY 989 ARQARLTLE----- 997
 Db 1086 AETVQPTMEDTNTTHPSGSEPOADTTQADPNSESVPSEETIEKVAENSPESETYAKORQ 1145
 QY 998 -----AKOVBPT--AKTOT----- 1009
 Db 1146 KATETTAQNDVAKKAPTEVANTQVIELAONGSETEETOAEFAROSEINSTEETVED 1205
 QY 1010 -----GEPKVSRR-----ABAAAF----- 1024
 Db 1206 DPTISEPKSRPSSISSNNINLAGTDTAKYETKTOEAPQVAFASFKOEHPBAKQ 1265
 QY 1025 -----POTLPDQ-----SLNMLAE----- 1039
 Db 1266 QROPKVOSQAKOBEITTOQAEPAFENVSTVNNYKEAQPOAKPTVAKEHTTASNEQKET 1325

QY 1040 -----KQABLPAETQSKAKT-----KKVSKSAV 1064
 Db 1326 AQPVANPEKTAENKAEENQSTETETDINHQEAHTAVASTEVTEENATPTIKENKTE 1385
 QY 1065 FSDPILLOQSFALE-----AALEYI----- 1084
 Db 1386 AQOVTETTTYSIENPVKNPENTTPATQSTQVNSEAVQSETATTEAVVQSQKVSABET 1445
 QY 1085 -----DAPQ--QSEKDR-----LAQBEAK-----QKQKDL 1109
 Db 1446 TYASQETTYDNGSGTTPQSPRSRRRBSAQNSIEPPELITENNAENPOGNDVATQVYRDL 1505
 QY 1110 ISRRNSGALSLSA-----TVNSMLSYODELRLFY-DQAGSAVNTNIAQDKERYDSAF 1163
 Db 1506 TETVNTNAVITDAMAQAQFVALNVGKAVSQISQLENNBGQYVWVNSGNTSMKENVSSQY 1565
 QY 1164 RAYQOQKTNARQVQKAL--ANGRGAVFHSNSDNTPEOVKNHATLITMSSGFAY-- 1219
 Db 1566 RHF--SKSAQTOQMDQITSSVOLGGVFTYVNSNNPD--KASSKNTLAQANLYSKTYMD 1623
 QY 1220 -QWEDLOFVNVGVTGISAKAEQSRKHRRKALNVGNVASYOFLQGLIOPYGVNRY 1278
 Db 1624 NEM--YIADLQIGNFQSNLOTHNAKFAHRTAQFGJLAKAENLGNFAVKTGYKYS 1660
 QY 1279 FIERENYQSEEVKTPSLAFNRNAGIRVDYTFPTDNISVKYFFVNTVDVSNANVT 1338
 Db 1681 YLSNANPALAKDKIKVNPISVYTAFAQVDLSYTY-HLGEFSITPILSARY-DANQSGSKI 1738
 QY 1339 TVNLTVLQOPFRGYWQKVGKALILHFOJSAFISGSGQLGQCVGVYKYG 1392
 Db 1739 NVDRIDPAVNVKQOQYNAGLKTKYHNVKLSLIGLITRAKQAEKQTAEXKLSF 1792

RESULT 3
 PCT-US04-07001-670
 ; Sequence 670 Application PC/US0407001
 ; GENERAL INFORMATION:
 ; APPLICANT: BAKULETZ et al.
 ; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
 ; FILE REFERENCE: 28335/38815A
 ; CURRENT APPLICATION NUMBER: PCT/US04/07001
 ; PRIORITY FILING DATE: 2004-03-05
 ; PRIOR APPLICATION NUMBER: US 60/453,134
 ; PRIOR FILING DATE: 2003-03-06
 ; NUMBER OF SEQ ID NOS: 771
 ; SOFTWARE: Patentia version 3.2
 ; SEQ ID NO 670
 ; LENGTH: 1542
 ; TYPE: PRT
 ; ORGANISM: H. influenzae
 ; PCT-US04-07001-670

Query Match 4.1%; Score 295.5; DB 1; Length 1542;
 Best Local Similarity 19.8%; Pred. No. 6.2e-06;
 Matches 299; Conservative 202; Mismatches 501; Indels 509; Gaps 74;
 QY 42 ENKGPVTGAGNIXKYVNOQVGTGNTKAPMIDFSVSRGVAALVENVQTVSAHVNG 101
 Db 219 KNEGVSIVNGSISL--AGQKITLIDILNPTITST-----AAPEEAL-----NLG 264
 QY 102 YTDVDFGABGNPDQHRFTYKLYKRNKYKCNLHPEYEDDYNHRLAKFYTEAAPIDMTSN 161
 Db 265 -----DIFAKGNINVRATTI--RKK-----GKLSAOSVSKSGSN 298
 QY 162 MNGSTVSDRTKYPEKVRIGSGRQFWRNDQKGD--VAGAVHYLTAGNTHNRQAGXGYS 219
 Db 299 IILSAKGGAEI-----SGVISAQNOQAKGKMTGKVTTLTGAVYIDLSGKGGSET 351
 QY 220 YLGADVAKAGEYGLPIAGSKDGSFMTYDAEKQKMLINGLIRGQNF--BKEGFPQ 277
 Db 352 YLGGDERGGRNG-----IQLAKTS-----LEKSGTINVSKEKGR 389

```

QY 278 LVKSYFDEIFERDCHTSLYTRAGNYTISGNDGSSITOKSG-----IPSE 326
DB 390 AI-----WKDIALIDGINKQSDIAKGGVETSGHLSIDSN 429
QY 327 I-----KILANMSLPLEKXD-KVHNPBYDGNPTYSRLNNGEYL----- 365
DB 430 AIYKTEWILDPDNYTIEAPSLSRADTDISEPFIODGFE-NSPKONADITILMETISN 486
QY 366 YEMOOK-----OGSLIFPSIDINGAGGLYFBSNFTVSPNSNQWQ 405
DB 489 ELONAKVYNTIARRKLTWNSSISIGSSHLIHSE-OGDGGVJIDEDIT-SEGGLTLIN 546
QY 406 GAG-IHVSNSYTFKAVNGVEHDLRSKIGTJHVAKGENKSSISVGGOKYILLEGQAD 464
DB 547 SGGWVDYHAKIT-----LGTFLINTRA-----GSSVAHEKG----- 577
QY 465 OGKQAFSEHIGLVSGRTVQLN-DDKQFDKFEYFPGGRILDLNGLSLTFKRIQNTD-- 521
DB 578 -GNABRNATDQITAGTITWKDKQ-----FRNNVSIINGBGLKFLIANQNF 627
QY 522 -----EGMAYVNHNTTOA-----NV-TTGNESI-----VLPNGNI 553
DB 628 CHKEDGEINISGLVITINOTYTKDAKWHASKOSYWNVSLSLTANDAKETFIKFVDSGNS 687
QY 554 NKL-DYKREJA-----YNGMFGEDKQ-KHNGRLNLTYKPT----- 587
DB 688 QULBARRRPAHYPHGTGKTNFNIGAKALFKLAKNMAATDPKELPITNNATNTG 747
QY 588 TEDRTL-----LISGGTNLKGDITGCKGLFFSGRPTPI-----AYKLAN 627
DB 748 SDDSVFEDIHANLTSRAASINMDSINITGDLFS--ITSINRSNAFEIKKDLINANT 805
QY 628 KMESEYEGIPQEIWDHDMINRTFKALNQKKGSAV--SRVSSISGNTYSNNANA 685
DB 806 SEFSLKPTKDLFENQYTGDAINST--RNLTILGKAVTLGGENSSNITGTTIAEANY 862
QY 686 TEGVVPNOQNTICTSDMT-----GLTCCQVUDTDTK-----YINSIP-KIOINSI 732
DB 863 TLQAYAD--NSIKKHKKTITLGNVSTSGNLSITSKYEVGDLAVLNKATRGKBTNDSL 920
QY 733 NL-----TDNATANY--KGLAKLNGVY-----LJNSQ-----F 760
DB 921 NITGFTNNGTADINIKGAVNIQDITNKGILNITTAQKQKQITLINNKKNMLNI 980
QY 761 TLSNNAFOI-----GWRIS-----DNSTAYDANL--N 788
DB 981 TNNNDTEIQIGNISQKESKMLTISDKVNITKOITIKAGVDEKSSSTADANLTIKT 1040
QY 789 GNHILDSAQFSLKNSHFSHOIOGDKGTVTLEN-----ATWTPSDTLQNTL 837
DB 1041 KMLVVEDNLISGFN--KAEITKAGSSDLITGNTNSADGNNAKVTINQYKDS---KIS 1094
QY 838 LNSTTILNSANSASSNN-----TPRRSLEPTEPTSAEHRFNTLTVNGKL 884
DB 1095 ANDHNVTILNSKVTSGNDNTGDSGNNAGLITIAKNEVYKNITS--KRTVNTIASEKL 1152
QY 885 SGO-----GTQFOTSLFGYKSDYKLSND--ABSDYTLISVRYTGGEPTLEQL 931
DB 1153 TTKADATIAATGNVEVTAKTGDIKGEVKSIGVNIYANGD-TLWNSVSGNAVTL-- 1208
QY 932 TVESKNOPLSDKLTLENDHYTAGA-RYKLYKNDGEF--RLHNPIDEGELANDVRA 989
DB 1209 -----TADKGLTTOSSSTISGTSVTSOSGDIGALISGNTSVYKATNDLI- 1256
QY 995 EOARTELEAKOVEPFAKTQGEPTVNSRPAAPAEPTLPDQSL-----LNLBPAK--- 1040
DB 1257 TKANSKEAK-----TGEANVVS--ANGILGIGISGNTVANTAGSLITIKOGAK 1304
QY 1041 -----QAEITAE-----TOKSAKTKYKVSRAVSPDLPDQSLFALEALAE- 1082
DB 1305 VDAENGATITAESKLTTOAGSTIISNNGOTTLTAKDGSIASSIDANAVNTLNTGTTLT 1364
QY 1083 -----VIDAQOSEKORLQOEAEKQKOKJLSIYSNSALSLS----- 1122

```

```

DB 1365 VQSSINANEGLTYINA-----QDATLNGDASGDRTEVNAVNSGSGSVTAVTSSVNT 1419
QY 1123 ---ATVNSMLSYOD-----ELDLFVD-----OASAVNTNITQDREYDSPAER 1164
DB 1420 GDLSTINGNITISNKGKTYVAKGAEIDYKTIQGVASAEVIEAKRLLEVKLSI--- 1476
QY 1465 AYQOQKTNLRQIYQOKALANGRIGAVFHSRSNDTPEQYKHAATLJWMSGFQYQWGL 1224
DB 1477 ---EERETLAKGVSAVFEVPEPMALITVTONFEPT---TRPSQVITSEKACFSSGD- 1528
QY 1225 QPGWVNGTIS 1235
DB 1529 --GAAVCTNVA 1537

RESULT 4
US-09-806-709-57
Sequence 57, Application US/09806709
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M.
APPLICANT: YANG, Yan-Ping
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
FILE REFERENCE: 1038-1138 MTS
CURRENT APPLICATION NUMBER: US/09/806/709
PRIOR APPLICATION NUMBER: PCT/CA99/00938
PRIOR FILING DATE: 1999-10-07
PRIOR APPLICATION NUMBER: 09/167,568
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 09/206,942
PRIOR FILING DATE: 1999-12-08
NUMBER OF SEQ ID NOS: 95
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 57
LENGTH: 1004
TYPE: PR1
ORGANISM: Haemophilus influenzae
US-09-806-709-57

Query Match 3.9%; Score 285.5; DB 5; Length 1004;
Best local Similarity 21.7%; Pred. No. 1.2e-07;
Matches 254; Conservative 169; Mismatches 438; Indels 309; Gaps 59;

QY 101 GYTDVDFGABGNPNDOHREFTYKIVK--RNN-----YK 130
DB 25 GREGDASNAKXNHNDKPTLNTTVENALKNTFPNITAKOKITVNSDINITKGAHLITLXS 84
QY 131 KXNLHPTEDDYHNRHLKHTVTAAPITKTSNMGSTYSRITYPE----- 175
DB 85 KXN-----KXSSVKLNGNITSTNGVLTLYSSGWDIHKNTITANTGY 126
QY 176 -PRATGSGROPVNRDOKGDOVA-----GAVHYLTAGNTNQGAGNGSYLGVDKAGE 230
DB 127 LNTLGGSVAFERAKNEKROVSESVYKAGCVITS-----GVGEGFR--NNVSLKCY 177
QY 231 YGPPIAGSKDSSGPMFTYDAEKOKMLINGILBGNPEKENGFOVLKSYFDEIFER 290
DB 178 GAGLFEVQKXNISN-----SVRENTL--KRRFDKMLNLSGVAVNSK--DVSGT 222
QY 291 DLHSLYLRAGNGYITISGNDGOSIT-OKSGIPEIKTITAMKSLPLEKOKVANNPRY 349
DB 223 KKHTRINRYWVNTTLNVAAGSSSFNLSIDASGISGMDODITVRGL-----NGITF 274
QY 350 DGNITYSPRLNNGEITYF-----MDOKGSS--LIPASDIN--QAGGLYFEGNFTVSPNS 400
DB 275 NBNENFN--LQGSINAHIKYSVNTREKSNALANGNISVLGGGYVNEHLN--ASSS 329
QY 401 NQTPQAGIHVNSNTVTKVNGVEHDLRSKIGKTLHVAQAG-----ENKGSISVG 452
DB 330 THTTSGALIN--SONFVNS-----GGSXLNKLKSGSTNPAFLIKNNLTINAT 374

```


Db 800 T-----KQVDSLAQ-DGSLAGISAAVNTINTTGLTVEGNIKATSGTL----- 845
 QY 955 VDAGALRYKLVKNDGEFRLHNPKEGCHLNDVRAAQERLLEAKQVEPACTQTGEPKY 1014
 Db 846 -----A-YNKAKLIDG-----TASGNRTVEMNANSGSGSY 876
 QY 1015 RSRRAAARPPDTLPDQSLNALFAKQELTAETOKSKAKTKYRSR--AVFSPDLDD 1072
 Db 877 TAKSSNV--NITGDLSTINGL-----NISENGR--TVRLRKEIDVKYIOP-- 921
 QY 1073 SLFPALEALVIDAPQOSEKCR-LAODEAKQKQKDLISRY--SNSAL-----SELAT 1124
 Db 922 ---GVASVEVIEAKVLEKVKDLSDEEBTLAKLGVAVFEPNNALTTINTQEFETR 978
 QY 1125 VSMALSYODELRLFPDQAQSAWNTNLAQD 1154
 Db 979 PSSQVIT-SEGRACFSSGNGAAYCTNVAAD 1007

RESULT 6

US-09-806-709-37
 ; Sequence 37, Application US/09836709
 ; GENERAL INFORMATION:
 ; APPLICANT: LOOSMORE, Sheena M.
 ; APPLICANT: YANG, Yan-Ping
 ; APPLICANT: KLEIN, Michel H.
 ; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
 ; FILE REFERENCE: 1038-1138 MIS
 ; CURRENT APPLICATION NUMBER: US/09/806,709
 ; CURRENT FILING DATE: 2001-04-04
 ; PRIOR APPLICATION NUMBER: PCT/CA99/00938
 ; PRIOR FILING DATE: 1999-10-07
 ; PRIOR APPLICATION NUMBER: 09/167,568
 ; PRIOR FILING DATE: 1998-10-07
 ; PRIOR APPLICATION NUMBER: 09/206,942
 ; PRIOR FILING DATE: 1999-12-08
 ; NUMBER OF SEQ ID NOS: 95
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 37
 ; LENGTH: 1222
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 ; US-09-806-709-37

Query Match 3.9%; Score 283.5; DB 5; Length 1222;

Best Local Similarity 21.4%; Pred. No. 2e-07;

Matches 268; Conservative 194; Mismatches 492; Indels 297; Gaps 65;

QY 40 PAENKG--KFTVGAQ-NIKVNRQGLVGTSMWKAPMIDPSVVSRYGAALVENQYVS- 95
 Db 135 FEDKPGLSMLTTAGTIAVNNKG-----FRDNTVLANGGGGSKYIETG 182
 QY 96 -----VAH-----NV-GYTDVDFGAEKNNPDQHRFTYKLVKNNYKNDLAPYEDYEN 143
 Db 183 NRDSNEETFRGRILNISGKVDILMCARCEMNRHWM-----GRSHWNTVLANSSESYEN 237
 QY 144 ERLKFTVTE---AAPIDMTSNMNGSTYSDETKKPEERYIGSRQPMKNDQKDYAA 199
 Db 238 VTTDSGSSASSPGAEPLMAQSLNGISENNDFV--NIAASSVNF-NIKPIVDKDTNG 294
 QY 200 YHVIAGNTHNRGAGNGSYLGGDYVRKAGEYGLPIAGSKDPSPMFIYDAEKQKMLI 259
 Db 295 NHTLEKGN-----ISVLDG-----GMSFTFINASSSYOT 324
 QY 260 NGILREGNPFEGKENGQYLYRSYFDELFEEDLHTS-LYTAANGVYITSGNDMGQSTIQ 319
 Db 325 YGVIIISQNFSA-SCGSSSLKFS-----EGSTYAA-FITKMDLILNATG---GNISLNG 373
 QY 320 KSGIPEIKITL-ANMSLPLKEDKXNPRYDGNL-----YSPLNNGELTYFMQKXG 373
 Db 374 VAGIDSNLKKSLAN-----KAITEGGNTLLAADKKPFIKGN---ITVKEG 418

QY 374 SLIFASDING--AGLIYEGNFTVSPNSNOTWOGACHVSNSTV-----TW 419
 Db 419 ANVLTASANYGNDXSAISIRGN--VTNKGNTLVGSAINIEKMLTVEGSAKFLANPVSF 476
 QY 420 KTVGVHNR-----LSKIG-----KGLTHVQKGN-----KQSTVGDQVILE 459
 Db 477 NYSGLFMDQGSNISLAKGALFKQIENTGSLNTTSSDSNHTTIKNTINRKGDLNT 536
 QY 460 QQADDQKQKQAFSEI---GLVSGR-GTVQLNDDKQPDYKRYF--GFRGGLDLN---GH 510
 Db 537 NNSDM-----TEIQGNISQKEGNLTISDQVNTIRIITKAGVNGDSSNATSA 589
 QY 511 SLTF--KAIQNTDGAIVYNTTQANVYITGESIVLPGANNIKLDYRKEIANGVF 568
 Db 590 NLITIKELKLTND---LINSFKAETITAKNSMLTIGDSKAGNDKAKTFFENVAD 645
 QY 569 GETDKKNGRNLNLYPFTTEDRTLLSGGNLKDITQTKKILFFSRPPTPHVHLAK 628
 Db 646 SKTSASDHRVTLNSKVE-TSGDSTEDGANNNG-LITIAKRVYNNNTISH--KTVNI 701
 QY 629 RNSMEGIPOGEIYWDHMINET-----FKAENFOIKGSAAVSNVSLIEGHWI--VSN 681
 Db 702 TASENVTTKAGT-----INATGSVEYIAKAGIDKIGIEGSGGNVITASGDTLANVN 755
 QY 682 NMAITFVPEQNTICTSDWTGULTTCQKDLITDKVINSIPKTOGNSINLTQANAN 741
 Db 756 ITGQVNTVAASAGAVTTGTSTINATGNANITTKGIEGNGVVS-ASGVNIT--ASGN 812
 QY 742 VGLAKLNG-NVTLNHSOFTLSNNAIOGNIRLSPNST--ATVDNANL----- 787
 Db 813 TLNVNINIGQNVYTIANS-----CAITTEGISTINATGDANITTOGNGINGV 861
 QY 788 ---NQNVELIDSAQSLKNSHPSHQ---IQGDKG--TVYVLENAWTMPSPDTLLQ---N 835
 Db 862 ESSSGSVTLIAEQ-TLAVNISGDTVITADKGLITQTSKINGKSVTSSGSGDIS 920
 QY 836 LTLNNTSTILNSAVSSANNTRRRLSEFTTTSKSHRENTLVNKGSLSGQCTQFTSS 895
 Db 921 GTISNTVSVASGSLITQASKIEAKTGBANVTSA-----TGIGIGTISG-NTVNVAN 974
 QY 896 LFGYKSDKLYKSNDAEGDYLISVNTSKE---PETLEQTLVESKQDQPSDKLEPTLEN 952
 Db 975 -----TDNLITIKQAR-----IKATGCAVTLTANIGTLITETSSDITSSNGQTLITAKD 1023
 QY 953 DHVDAKALRYKLVKNDGEFRLHNPKEGCHLNDVRAAQERLLEAKQVE----- 1002
 Db 1024 SSI-AGSI-----NANVLTINTGTLTVAQSKIEAASGTLVINAKDAQLOGAASGDHT 1076
 QY 1003 -PLACTQYGEPRKRSRAARAFDPTLPDQSLNALFAKQELTAETOKSKAKTKYRSK 1061
 Db 1077 VVNAINANGSGSVIATTSRNVNITGDILITINGLNTI--SKNKNVTLKGVIEDVYIOP- 1134
 QY 1062 RAVFEDPLDQSLFALEALVIDAPQOSEKCR-LAODEAKQKQKDLISRY--SNSAL 1118
 Db 1135 -----GIASVNEVIEAKKALEKVDLSDEERTTLAKGVAVFAPNNAI 1180
 QY 1119 SELATVNSMLSYODE---LDLFPVDAQ-----SAVNTNLAQDKR 1156
 Db 1181 -----TINT-----QNEFTTRPLSQVITISEGKVCFLINGCATCTINADIER 1222

RESULT 7

US-09-806-709-34
 ; Sequence 34, Application US/09806709
 ; GENERAL INFORMATION:
 ; APPLICANT: LOOSMORE, Sheena M.
 ; APPLICANT: YANG, Yan-Ping
 ; APPLICANT: KLEIN, Michel H.
 ; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
 ; FILE REFERENCE: 1038-1138 MIS
 ; CURRENT APPLICATION NUMBER: US/09/806,709
 ; CURRENT FILING DATE: 2001-04-04

; PRIOR APPLICATION NUMBER: PCT/CA99/00938
 ; PRIOR FILING DATE: 1999-10-07
 ; PRIOR APPLICATION NUMBER: 09/167,568
 ; PRIOR FILING DATE: 1998-10-07
 ; PRIOR APPLICATION NUMBER: 09/206,942
 ; PRIOR FILING DATE: 1999-12-08
 ; NUMBER OF SEQ ID NOS: 95
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 34
 ; LENGTH: 1228
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 US-09-806-709-34

Query Match 3.9%; Score 283.5; DB 5; Length 1228;
 Best Local Similarity 21.4%; Pred. No. 2e-07;

Matches 268; Conservative 194; Mismatches 492; Indels 297; Gaps 65;

QY 40 EAENKG-KFTVGAQ-NIKVYKQQLVSTSTKAPVDPVSYSNGVALVYNOYTVS- 95
 DB 141 FEDKPELSMLITAKTIAVNNKQ-----FREDVTLNGGGGSEKTLITG 188
 QY 96 -----VAH-----NV-GYTDVDFAGSNNPDQREFTYKIVKNNYKKNLHPYEDDYN 143
 DB 189 NRDSNFEHFRGRNLISGVKIDILMOARQENNRHW-----GRSHMVTILVNSNSYFN 243
 QY 144 PRLHFTVE-----AAPIDTSMNNSITSDRCKPERRVIGSGQFNRNDQDQVACA 199
 DB 244 VITDSSSASSPGAGLNMOSGLNLSFNDTVF--NIASSAVNF-NIKRPIDVTVNG 300
 QY 200 YHVLPAKTNMORGANGVSYLGGVRAKAGEYGP-LPIAGSKDSSGPFYIDAEKQKWL 259
 DB 301 NHTLPKKN-----ISVLGG-----GMSTF-PNASSSNVOT 330
 QY 260 NGILFEGNPFEGKENGFOVRSYSYDEIFERDLHSLYTRAGNGVTTISGNDGQSITQ 319
 DB 331 YGVILISQNFSA-SGSSSLFKFS-----EGSTHAA-FTIKDILNLTG--GNLSINQ 379
 QY 320 KSGIPSEIKITL-ANMSLPLKKKKKQVNRPRVDGPN-----YSPLAKSGILYPMQKQ 373
 DB 383 VAGIDSEIKKSLAN-----KNITFEBSGNTLAADKKPEIKKN-----ITVEEG 424
 QY 374 SLIFASDING--AGGLYFEGNFTVSPNSQVQAGLHVENSTV-----TW 419
 DB 425 AAVTLRSANYGNDKSLISRN--VTMKNGLTVYGSANINERULTYHSGSKPLAIPNYSF 482
 QY 420 KNGVEHDR-----ISKIG-----KGLHVOAKGEN-----KGSISVGDQVILE 459
 DB 483 NVSGLEPNQGSNISIAMKGAIFPDIENTGSLNITKSDSNHHTTIKKNITNRKGLDMLIT 542
 QY 460 QOADOQNKQAFSEI---GLVSGR-GTVJOLNDKQPDNDKPYF--GPRGRDLN---GH 510
 DB 543 NNGDN-----TEIQGNISQKEGWLITSSDKNVTERTTITAGVAGNDSDBATSA 595
 QY 511 SLTF--KRIQNTDEGATVNNNTQAAVTTIGNESIVLPNGNNINLDPREKILVYNSWF 568
 DB 596 NLTIKTELKLTND-----LNTSGFNKKEITAKONSMLTIGNSDAGNTAKVTPSNVQD 651
 QY 569 GDTCKNENGRMLLYKPTEDRLLISGNTLAKGDIQYQKLFSSGAPTEPHAVNHLNK 628
 DB 652 SKLISHENVTLSKYE--TSGDITDSTEDGNNNTG-LITTKAVTVNNNITSH--KTVNI 707
 QY 629 RWSIEGIGPGEIWDHMINRT-----FRAENFOIKGSGAVSVRNVSSIEGNTV--VSN 681
 DB 708 TASENVYTKAGT-----INATTSVEVTKAGDKGKIGNSGNNTTASGDTLVNSN 761
 QY 682 NAAATFGVBNQNTCTIRSDWTGLTTCQVLDITGVKINSIPTQINGISILTNATNATN 741
 DB 762 ITGQNTVTAASGAVTTTKSTINATTGNAANTTTGSEINSEYKS--ASGNVIT--ASGN 818
 QY 742 VKGLAKNG--NTLNNHSGFTLSNATQIGNRLSDNST--ACTDNANT----- 787
 DB 819 TLNVSNTGQNTVTVANS-----GALTTEGSTINATTGDANITTOGNGKV 867

QY 788 ---NGNVHLTDSAQFELKNSHPSHQ---IQDKG--TIVTLENAWTMPEDTTLQ-----N 835
 DB 868 SSSSSVTLIATGQ-TLAVNGISGDTYITADGKGLTQTSKINGKSTISSQSGDLS 926
 QY 836 LILNNSITTLNLSAVSASSNTPRRRSLETFTTPTSAHRPNTLVNKLSGQGTQFTSS 895
 DB 927 GTISGNTVSVASGSLTTOAGSKTEAKTGEANTSA-----TGITGGTISG--NTVAVNAN 980
 QY 896 LFGYSDKGLKISNDAGDYILSVNTKE---PELEQLVLESKQNPISDQKXTLEEN 952
 DB 981 -----TNLTIYKQAR-----IKATGAVTLTARGTLITETSDITSSNGQITVLAKD 1029
 QY 953 DWDAGALRYKLVNDGFERLNFPEIKQELANDLVAAEQARTLEAKQVE----- 1002
 DB 1030 SSI-AGSI-----NAAVNTLNTGLITLVASGKTEAASGLVIMADQDLGAASGHT 1082
 QY 1003 -PTKCTQGEPRKRSRARAAPDTLPDQSLNALMAKQELTAETQSKAKTVKVSX 1061
 DB 1083 VMAITNANGSGVLTATSSRVNTIGDLITINGLNTI-SKNGKNTVLKQVEIDVRYIQP- 1140
 QY 1062 RAVFSDPLDQSLFALEALAVIDAPQOSKDR-LAOEAKKQKQKDLISRYS--NSAL 1118
 DB 1141 -----GLASVNEVIAKRALBKVDLSDEKETLAKGVSAVPAEPNNAI 1186
 QY 1119 SELSATVNSMLSVQDE-----LDELTVDAQ-----SAWNTNIAQDKR 1156
 DB 1187 -----TLNT-----QNEFTTRPLSGVTTISGKVCFLINGANGITCINIADIER 1228

RESULT 8

US-09-806-709-67

Sequence 67, Application US/09806709

GENERAL INFORMATION:

APPLICANT: LOOSMORE, Sheena M.

APPLICANT: YANG, Yan-Ping

APPLICANT: KLEIN, Michel H.

TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH

TITLE OF INVENTION: MOLECULAR WEIGHT PROTEINS

FILE REFERENCE: 1038-1138 MIS

CURRENT APPLICATION NUMBER: US/09/806,709

CURRENT FILING DATE: 2001-04-04

PRIOR APPLICATION NUMBER: PCT/CA99/00938

PRIOR FILING DATE: 1999-10-07

PRIOR APPLICATION NUMBER: 09/167,568

PRIOR FILING DATE: 1998-10-07

PRIOR APPLICATION NUMBER: 09/206,942

PRIOR FILING DATE: 1999-12-08

NUMBER OF SEQ ID NOS: 95

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 67

LENGTH: 1536

TYPE: PRT

ORGANISM: Haemophilus influenzae

US-09-806-709-67

Query Match 3.9%; Score 283.5; DB 5; Length 1536;

Best Local Similarity 21.0%; Pred. No. 2.7e-07;

Matches 244; Conservative 162; Mismatches 400; Indels 357; Gaps 56;

QY 42 ENKGRFTVGAQINIVNKKQQLVGSSTKAPMIDFSVNSKNGVALVNOYTVSVANNQ 101
 DB 219 KNEGVSIVNGSISLT--AGQKITISDIINFTIYTS-----AAPNE-----AVNIG 264
 QY 102 YTVDFGAEAGNPPQHRFTYKIVKNNYKKNLHPYEDDYNPRLHKFTVTEAAPIDTNS 161
 DB 265 -----DIKAKGINVRAAT-----RNO-----GKLSADSVSKQSGN 298
 QY 162 KNGSTYSDBTKPERRVIGSGQFNRNDQDQD--VAGAYHVLPAKTNMORGANGVYS 219
 DB 299 IVLSAKEGHAB-----IG-GVISAQNGQARGKGLMTGDKVTLTKTGAVIDLISGKGGGT 351
 QY 220 YGQGVRAKAGEYGPLPIAGSKDSSGPFYIDAEKQ-----KW-----LING-ILREKNPFE 270

Db 352 YIGDERGEGKGIOLAKKTSLKSGSTINVSKEKGRALVWGDALIDENIMAGSSDI 411
 Qy 271 GSENGFOLVRSYEDLIFERD-----LHIS---LYTRAGN 302
 Db 412 AKTGSGF-VETSGHD-LFKCNALVDAKEMLDPONVSNINAGRSNISEDEVTSGN 468
 Qy 303 GYTTISGNDNGGSILOKSGLPSEIKITILANMSP--LKECXVH--NPRDGNIT- 355
 Db 469 SASFPYERKE-----KTLINLTLEILAKGTFINITAMOR-----ITVN 508
 Qy 356 -SPRLNGETLLYPMQOKOSLIPASDINOGAGLYFEAGNFTVSPNSNOTMOGAGIHVSEN 414
 Db 509 SSILNWSGLTIWSEBSG--GGEVINDITGDDT-----KAMLTISG 552
 Qy 415 SYTWKVVGVYEHDLRSKIGKTHVQAKSGMKS-SVGDGVYILFQADDOGNKAFSEI 474
 Db 553 GWVD-----VEKNSLGAQGNINITAK-----QDIAFEKGSNO----- 585
 Qy 475 GLVSGRGTVQNDKQFDTKPYFGFGRGLDINGH-----SLTEPKR-----IQNDEGAM 525
 Db 586 -VITGQGTITSGNQK-----GFRPNVSLNIGTSGLOFTIKRINKKATINAFESTL 635
 Qy 526 IYVHNTQAAVNTTIGNSIYLPNGNNINKLDYRKELIAYNGMGETDKN-----KXNGR 579
 Db 636 -----NISGKNISWVLP-----KHSGYDKFKGTVMWLSLVNSSEGE 675
 Qy 580 INLIYKPTDRITLISGTVLXGDTIQKGLFSGSGFPEVHYNHAKRMSBEVGPQ 639
 Db 676 ENL-----TIDSR-----GDSAGTLTO-----YXN-----ANGISFN 703
 Qy 640 EIVWDHDMINRFEKAB-----NFOIK-----GSAVAVSRVSSIEGNMTVSNANNAFPVAV 691
 Db 704 R-----DITFENBARAVNPDIAPPIGINKYSILNIVASPNGNISVSGGSGVDTFLA 755
 Qy 692 NQOVT-----ICRSWITGLTQCKV-----DCTDKVINSIPYQIUNG 731
 Db 756 SSSNVQTPGVVINSKYFVNSYSSLSFKTSGSTKIGFSEIKEDLTINATGNTLLQYEGT 815
 Qy 732 INLTDN-----ATANV-----KGLAKINGVTLTNHSGFTLSRNATQION----- 771
 Db 816 DGMIGKIVAKKNIIFBEGNITFGSKAVTEIEGVITINNA-----NVTLIGSDEDMH 869
 Qy 772 -----ILSSTAVAVDANLNGVH/LTSAOFSKNSHFSHOIOG--DKG-TT 817
 Db 870 QKPLTKKQVINSNGLIAGNIVNLAGNLVRESNANFK-AIYNFTVNGGLPUNKGNSN 928
 Qy 818 VTLNNAVTMPBDITLQNLTL-NKSTITLNSAYASSNTPRRRSLETTETTSAB----- 872
 Db 929 ISIAKGAARFKOINDSKNLSITTSNSTYRTIISGNITKNAGDNIIVEGSDTVMQIGD 988
 Qy 873 -HRENTLVNG-KLSGQGTFCPTSLSFGVSD-----KXKLSND----- 909
 Db 989 VQKREGNLTISDKINTTKOITIKAGVGENSDATNNANLITKTKELKLTODLNIISGF 1048
 Qy 910 -----AESDYILSVNRTGKEPET-LEQTLVSEKNOPLSKLFTLENDHVDAGALY 962
 Db 1049 NKAELTAKGSDLTIGTNSADQVNAKVTYVQKOSKISADGHVTLHSVETSSNNN 1108
 Qy 963 KLVQNDGEFRLHPIKEQHLNDIYPAFCARITLAKQVEPFAKTQGESEPRVSRPAPA 1022
 Db 1109 TEDSSDNNAGLTIDAKVAVVNNNITSHKAV--SISATSGEITTKGT-----TIVATIG 1160
 Qy 1023 AFEDTLPOSLNANLEAKQALY 1045
 Db 1161 NVEITAGTSGITIGISSGSVY 1183

RESULT 9
US-09-806-709-71

Sequence 71, Application US/09806709
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M.

APPLICANT: YANG, Yan-Ping
 APPLICANT: KLEIN, Michel H.
 TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
 TITLE OF INVENTION: MOLECULAR WEIGHT PROTEINS
 FILE REFERENCE: 1038-1138 MIS
 CURRENT APPLICATION NUMBER: US/09/806, 709
 PRIOR FILING DATE: 2001-04-04
 PRIOR APPLICATION NUMBER: PCT/CA99/00938
 PRIOR FILING DATE: 1999-10-07
 PRIOR FILING DATE: 1998-10-07
 PRIOR FILING DATE: 1998-10-07
 PRIOR FILING DATE: 1999-12-08
 NUMBER OF SEQ ID NOS: 95
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 71
 LENGTH: 1477
 TYPE: ERT
 ORGANISM: Haemophilus influenzae
 US-09-806-709-71

Query Match 3.8%; Score 280; DB 5; Length 1477;
 Best Local Similarity 23.2%; Pred. No. 4e-07;
 Matches 230; Conservative 151; Mismatches 402; Indels 358; Gaps 49;

Qy 42 ENKGFYVGAONIKVYKQOLVTSMTKAPMIDFSVSRNGVAALVENQYIVSAHNG 101
 Db 219 KKEGVIVNGGSIIL--AQKILISDLINPTITSI-----AAPENE-----AVNIG 264
 Qy 102 YVDVPGAEVNPQPHFTYKIVANNYKNDLHPYEDDINPRLHFEVTAAPITMSTN 161
 Db 265 ---DIFAKGINVRAAT--RNO-----GKLSADVSKXKSGN 298
 Qy 162 MNGSTYSRTPYPERVIRIGSRQFWRNDQKGDQ--VAGAHVYLACNTNQGANGIS 219
 Db 299 IYLSAKGEAE--IG-GTISNOQAKGKGMITGDKYTLTGAVIYLSGKEGET 351
 Qy 220 YIGDVRKAGEYGLPIAGSKDGSMPFIYDAKO--KW-----LINS-ILREANPE 270
 Db 352 YIGDERGEGKGIOLAKKTSLKSGSTINVSKEKGRALVWGDALIDENIMAGSSDI 411
 Qy 271 GSENGFOLVRSYEDLIFERD/LHISLYTRAGNYTISND-----NQOG 315
 Db 412 AKTGSGFVETSGHYSIDSNALVTKEWLDDPV-TTAEBDPLNNTGINDFFPTGTGEA 470
 Qy 316 SITOKSGIPSEIKITILANSLPIKEDKVANPRYDGPVYSPRLNGETLYMDOKQSL 375
 Db 471 SDPKKN--SEKTTITNVTI-----SNYKXAMTWNITASRK--L 506
 Qy 376 IPASDINQA-----GALYFEGNFTVSPNSNOTMOGAGIHVSENSVTWTVN 422
 Db 507 TVNSINIGNSHLILHSGQGGGVQIDEDITSGKGLTIYSGWVDVHKUL----- 560
 Qy 422 GVHEDRLSKIGKGLTHVQAKENGSGISYGDQKVLLEQADDOGNKQASSEIGVSGRT 482
 Db 561 -----LDQFNLNTA--ASVAFEG--NNKADANAK--TVAOGT 595
 Qy 483 VQNDKQFDTKPYFGFGRGLDINCHSLTFKRLQNTDEGAMIVN-----HNTTQAN 536
 Db 596 VITIGBEK-----DEPRANVSLNG-----TGKGLNIISSVNNLTHLSGIN 637
 Qy 537 VT--ITGN-----SIVLPNGNNINKLDYRKE----- 561
 Db 638 ISGNITLNOTRKNITSYQTSYSHSHNVSAALMFTGAMFTFIKYSNSKGLITQYSSA 697
 Qy 562 -IAYGWFGEYDKN-KXNGRNLVYKPTEDT-----LLSGCNLKGDTIOT 608
 Db 698 GVNNSJNGNMSFNLKBAKYNFKLKENNMTSAPLPIPLANTATG-- 747
 Qy 609 KGKLF-----SGPTEPHAYNHLN-----KRWSEMEGIPGEIYWD----- 644
 Db 748 -GSVFEDIVAHNSGGAELKMSSEINISNGAFILNSVVRDDAFKINDLTINATNSNFS 806

QY 645 -----HMINRTPKAENPQKGSAAVY--SRVSSIEGNTVSNANATPGV 689
 Db 807 LQOTKDFYDGAFAINASTY--NISIIAGVTLIGSSSSSITNTITTEKAAVTLA 863
 QY 690 --VPMQONTICRSPWTLTTCQVDLTDTKVINSIPKQINGSI--MLTMAATANYKH 745
 Db 864 NNANQON-----IRDYIKGSLVNGSLSTGENADIKMILTISEATPKG 912
 QY 746 AK--LNGVTLTNHSGFTLSNNAQIGNI-----RLSDNSTATVDMANLNG 789
 Db 913 TROTLNITGNFTNMTAFINITOGVVKLGNTGDMINTTHAKRQUSIIIGDIIKKG 972
 QY 790 NHTLTSQAQSLKNSHSPHOIGDKGTTLLENATWTMPSD--TTLOKLTNNGSTIAMS 847
 Db 973 SLMTDS-----NMDAIDQIGN--ISOKEGNLTISDK.NITKOITIKKIDGDS 1022
 QY 848 AYGAASNTTPRRASLETETTPSAHRENYTLTNGKLSGGTFOPTSSIFPKSDKLKS 907
 Db 1023 ISDATSNANULTIKTKETLTDLSISGFKAKITAK-----DGRDLITG 1066
 QY 908 NDAEGDYI.SVANTGKEPETLEQLLVESK-----DNQPLSDKLTLENDHDAGALRY 962
 Db 1067 NSWDG-----NESAIAKVTTFNNKOSKLSADGRVTLNKRVTSSNGSRS----- 1114
 QY 963 KLVKNDSEFLNPKQELHNDLVRAE-----QAEPTLEAKQVEPTAKTONG 1010
 Db 1115 --NSDNDTGLITAKNVEYKDIISLKTNTVITASKEVYTTAGSTIATKAKSITTKIG 1171
 QY 1011 E 1011
 Db 1172 D 1172

RESULT 10

US-09-806-709-53

/ Sequence 53, Application US/09806709
 / GENERAL INFORMATION:
 / APPLICANT: LOOSMORE, Sheena M.
 / APPLICANT: YANG, Yan-ping
 / APPLICANT: KLEIN, Michel H
 / TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
 / FILE REFERENCE: 1038-1138 MIS
 / CURRENT APPLICATION NUMBER: US/09/806,709
 / PRIOR FILING DATE: 2001-04-04
 / PRIOR APPLICATION NUMBER: PCT/CA99/00938
 / PRIOR FILING DATE: 1999-10-07
 / PRIOR APPLICATION NUMBER: 09/167,568
 / PRIOR FILING DATE: 1998-10-07
 / PRIOR APPLICATION NUMBER: 09/206,942
 / PRIOR FILING DATE: 1999-12-08
 / NUMBER OF SEQ ID NOS: 95
 / SOFTWARE: PatentIn ver. 2.1
 / SEQ ID NO 53
 / LENGTH: 977
 / TYPE: PRN
 / ORGANISM: Haemophilus influenzae
 US-09-806-709-53

Query Match

Best Local Similarity 3.8%, Score 277, DB 5, Length 977,
 Matches 241; Conservative 164; Mismatches 436; Indels 308; Gaps 52;

QY 123 IYKRNVYKONLHPYEDDYHNPRLHKFVTAAIDMTSNMGSYSRTYPERVRISSG 182
 Db 17 VYRGSIEKKA-PTTITTHASISIEGIMKGAVINISANKVYTTDINVYGAALLHSE 75
 QY 183 RQFMKNDQKGDQVAGAVHYLTJAN-----THQAGANGY-SYIAGGVAKAGEY 231
 Db 76 RD-----GVEINGNITSEKGNLTITAGSVVYHKNILTSGFNLTSSD----- 120
 QY 232 GBLPLASKGDSSSEKFTIYDAKOKMLINCILLRENGPEKGENGFQVLYKSYFDEIFERD 291

Db 121 ----IAFEKGNLIT-----ITAQGNITSNKD-GKOL-----RL 148
 QY 292 LHTSLY-TPAQ-----NGYITISDNDGQSSIIQKSGISEKLTITAMSLPKEXD 342
 Db 149 NNVSLNGTAGLFIANONNFTNHSIGALINSVLT-----INQTKKKA 193
 QY 343 KYHNPRIYDGP-NIYSPRLNNGETIYFMDQKGSILIPASDINOAGALYFEGNFTVSPNSN 401
 Db 194 KAMNTSYDSYWNVSTLTLSNDAKFTFK-----YVDSNHSNTSSDS 234
 QY 402 QTWQAGIHSVSNSTYTWKY-NGVEHRLSTIKGCTLHVQKGENGSISV-GDKVTL 458
 Db 235 RSFAGVKEH-GKONNEKFEWIGNNAKAEFRILKPNKETTTPRPLPQIPLNSISVGGGVFE 293
 QY 459 EGOAD--DOGNKQAFSEIGLYSAR-----GTQVANDQOPD-----TD 494
 Db 294 DLYANLWKGTLEKMSINVSSSNLTLNSHVKNYAFININDTITNATNSFNLRQTS 353
 QY 495 KFYFGFR-----GGRDLNGH-----SLTEKRIQ-----TDEGAMIV 527
 Db 354 SFERNGYRNMAINSTANISILGONVTLGGQSSSSIMENIIKRAANTLEADNSHSDV 413
 QY 528 NNTTQAANTTIGNSIYLPKANNINKLDYKELAVNGWFGETDKRKNAGRLNLIYKPT 587
 Db 414 KORTINLGNLTVEGNSLIGENANINGLSIEKEAIFG--XTKDSLITGNFTNNGT 469
 QY 588 TEDRTLLSGGTLLKGDITOTKGLFPSPGRPHAVYHNLKRMSEMGIPQGIYWDWD 647
 Db 470 AE--INISQVSLGDIIT-NDGL-----NITTHAK-----SGQKSIIRGDIINKGN 514
 QY 648 INRTPKAENPQKGSAAVSNVSSIEGMWTVSN-ANATPGV-----VNDONTICTRS 701
 Db 515 LNTIDNNSNALEIRIG--NISOKEGNLTISDKNTITQITIKAGVDESSSSSIA 569
 QY 702 DMTGTLTCKVDLTDTKVINS-PTQINGSN-----LITNATH-----NYGLAK 747
 Db 576 DANLTITKELTIDNLNISEFNFAE-TAKDSDLIIGKASDNSNAKQVTEPKDKSKI 629
 QY 748 LNSGVTLTNHSOFTLSNNAQIGNIRISDNSTATVDNANLGNVHLTDSAGSLKNSHF 806
 Db 630 SAGHNVTNLSKRVTSN-----SDSGTG--NSDUNNIGLITISKQVTVNSIT 676
 QY 807 SHQ-----IGDKGTTVLENAWTMPSDTLLOMTLNNSTITLNSAYSNASNT 856
 Db 677 SHKVNISASBEGITTKAGITIVATTGSVEYTAKGDISGITSKIVSV--TASGDL 732
 QY 857 PRSSLETTPTPSAHEHNTLTWNGK-LSGGTFQPTSSIFPKSDKLKSNDAEGDYI 915
 Db 733 TVRKAATISATGAA-----TLTATGNTLTTEAGSSITST-----KGVDLSAQ-DGSLA 781
 QY 916 LSVNTEKPEPTLEOLTVESKDNQPLSDKLTLENDHDAGALRYKLVKNDDEFLHN 975
 Db 782 GOISAAVNTLNTTGLTLTVBSNLIKATSGTL-----AIYADAKLID-- 823
 QY 976 PIKQELHNDLVRAHQARTLEAKQVEPTAKTQGEKVSRRAPAPAPPTLPDQSLN 1035
 Db 824 -----TASGRTEVATWASGSQSVYAKTSSNV--NITGDLSTIN 861
 QY 1036 ALEKQAELEAETOKSAKTKVYRSK-AVRSPELDQSLTELEBALEYDAAQSEKD 1093
 Db 862 GL-----NISMGN--TVRLKGEIDVKTID-----GVASVEEVLBAKRVLEKY 906
 QY 1094 R-LAQBPAKOKOKDILSY--SNSAL-----SELSATVNSMLVQDELRLRFPVDAQS 1145
 Db 907 KDLSEBERTLAKGVSAVFEPNNALITMONEFTTRSSQVII--SEGACSSGNGA 965
 QY 1146 AYVNTIAQD 1154
 Db 966 AVCTNVAAD 974

RESULT 11

US-09-806-709-51

```

; Sequence 51, Application US/09806709
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-ping
; APPLICANT: KLEIN, Midee H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; FILE REFERENCE: 1038-1138 MIS
; CURRENT APPLICATION NUMBER: US/09/806,709
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: PCT/CA93/00938
; PRIOR FILING DATE: 1999-10-07
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 09/206,942
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 51
; LENGTH: 983
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; US-09-806-709-51

```

```

Query Match      3.8%; Score 277; DB 5; Length 983;
Best Local Similarity 21.0%; Pred. No. 3,4e-07;
Matches 241; Conservative 164; Mismatches 436; Indels 308; Gaps 52;

```

```

QY 123 TVKNNYKDNLPYEDDYHNPRLKFTVTEAPFIMTSMNNGSTYSDRKTPERYRIGSG 182
DB 23 VVRSDSTKKNA-PKTYTHAGSIEJOSIKKGAANYISATKRVVTDIVVYGAALTLSHE 81
QY 183 RQFWRNDQKQCVAGAYHYLTAGN-----THQKAGNGY-STYGGDPRKAGXEY 231
DB 82 RD-----GVINGNITSEKGNLTIRAGSWVYKNTTJGEGLTATSGD----- 126
QY 222 GPDLIAGSKGSGSPMFTYDAEKQKMLINGILREGNPEKENGFPOLVYKSYFDEFERD 291
DB 127 ---IAFEKGNMLT-----IAQGNLTISNKD-GKQL-----RL 154
QY 222 LHTSLY-TRAG-----NGVTTISGNNGGSGITQKSGIPSEIKITLANKSLPLEKED 342
DB 155 NNVSLNGTGAGLNFANONNFTNHSALNISGVVT-----INQTKKNA 199
QY 343 KYHNPRIYDGP-NIYSPRLNNGETLYFMDQKQSLIFASDINGAQAGLYPEGAFVSPNSK 401
DB 200 KAWMTSYDSYVWVSTLTLSDNAKFTPIK-----YDSNHSCTMSOS 240
QY 402 CTWQAGAGIHVSENSTVTKV--NGVEHRLSKIGKTLHVOAKENKGSISY-GDGKVTIL 458
DB 241 RSPAGVYFH-GKCNEMKFNIGNNAKAFPLKFNKPTPNRPLPIQPLANSIVTGGSVFP 299
QY 459 EQQAD--DOGNQASSELGIVSGR-----GVQALNDKQD-----TD 494
DB 300 DTYANLKGKTELKMDISVSGSNLTLSHYRKYNAFELINKDILTNTNSFNLRQTSN 359
QY 495 KEYFGR-----GGRLDINGH-----SLTFEKLQV-----TDEGAMTY 527
DB 360 SFRGVRNNAINSTHNISLIGAVTLTGCGNSSSSIMKNIITIKAAVTELEADNSHNSDV 419
QY 528 NHATTCAAVTTITGESVLPNGANNINKLQYKKEIAYNGVFGEDTKNKRGNRLMLYKPE 587
DB 420 KDRINLNTLVTEGNTLSIGENANINGNISLEKZA-FKG--KTKDSLNTIGFTNNGT 475
QY 588 TEDRTLLSGGNTLKGDLTQTKGLFFSGRPTPHAYNHLKMKRSEMEGIPGEGELVWHDHW 647
DB 476 AE---INISQVSLAGDIT-NDGKL---NITTHAK-----SQQKSLIRGDLINKQCN 523
QY 648 INTFEAFENFOIKGSAVSRVNSIEGNMTVSN--ANATFGV-----VPMOQNTICTRS 701
DB 521 LNTDNNNSVAHEIGG-----NISCKEKNLTISDDKNITKQITIKAYVDESSSSSTAS 575
QY 702 DWGTLTCKQKVDLTDITVINSIEKTOJNGSIN-----LTDNATA-----NVKGLAK 747

```

```

DB 576 DAULTTKTELTLDNINISGENKKEITAKONSDLIIGKASSDNGSNAQVTFEDYKXSKI 635
QY 748 LQNGVTLNHSQPTLSNNAQOINIRLSDNSTATVDNANNGVHTLSAQOFSKNSHF- 806
DB 636 SAGNHVTLANSVETSN-----SDGSGT---NGSDNNIGLITSAQVTVNSNLT 682
QY 807 SHQ-----IQGKGTTLTLENATWMPSDTLQMLTNNSTITLANSASASNNMT 856
DB 683 SHKTWNISASEGGITTAGTTINATGASVEVYAKTGDISITISGKTVSV---TASIGDL 738
QY 857 PRBSLETTPTSAHERPVLTVNCK-LSQGTFOPTSSLPKYSDLKLSNABEDTI 915
DB 739 TVKKAATLSAEGGA-----TLAAGMTLTTEAGSSITST---KQGVDSLAC-DQSLA 787
QY 916 LSVANTGKEPELTBLQTLTVESKNOPLSDKLTLENDYDAGALRYLVKNOGDEPRLHN 975
DB 788 GQISAANYTLMTGTTLTVGGSNIKATSGTL-----ALNADAKLDO----- 829
QY 976 PIKEDELHNDLVRAEQARTLEAKQVEPTAKTQTEPKYRSRARAAPPTLTPDOSLNL 1035
DB 830 -----TASGNRTVNAATNAGSGSVTAKTSNV---NITGDLSTIN 867
QY 1036 ALFAKQBELTETQKSAKTKKYSKR--AVPSDPLDGLFALMALETVIDAPQOSEPD 1093
DB 868 GL-----NISSEKRN---TVRLRKEIDVXIOP-----GVASVEVTEARVLEKY 912
QY 1094 R-LAQEAEKQKOKOLISRY--SNSAL-----SELSATVNSMLSVODELRLFPVQAO 1145
DB 913 KDLSDERETLAKJGVSAVRFEPNNAITNTQNEFTPTSPSSQVIT-SEKACACSSGNA 971
QY 1146 AWTNTIAD 1154
DB 972 AVCTNVAAD 980

```

```

RESULT 12
PCT-US04-07001-693
; Sequence 693, Application PC/TUS0407001
; GENERAL INFORMATION:
; APPLICANT: BAKALET et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONYPEABE H. INFLUENZAE
; FILE REFERENCE: 2835/538115A
; CURRENT APPLICATION NUMBER: PCT/US04/07001
; PRIOR FILING DATE: 2004-03-05
; PRIOR APPLICATION NUMBER: US 60/453,134
; PRIOR FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 771
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 693
; LENGTH: 1381
; TYPE: PRT
; ORGANISM: H. influenzae
PCT-US04-07001-693

```

```

Query Match      3.8%; Score 275; DB 1; Length 1381;
Best Local Similarity 19.1%; Pred. No. 6.8e-07;
Matches 274; Conservative 192; Mismatches 484; Indels 484; Gaps 62;

```

```

QY 42 ENKGFVGAQNTIKVYKQGLVGTSMYKAPM-IDFSVSRNGVAALVENQYISVAHVNG 101
DB 108 KNEGVISVNGSISL--AGQKITISDIINPTTYSI-----AAPEHEAT-----NLG 153
QY 102 YTDVDFASANNDOGRFTYKIVKANNYKKNLHPYEDDYNPRLKFTTEAPIDMTSN 161
DB 154 ---DIFAKGNIN-----VPAANIRNG-----KLSADSVSDKSGN 187
QY 162 MNGSTYSDRTKYPERVATGSGQFWRNDQKQD--VAGAYHYLTAGNTHNRGAGNYS 219
DB 188 IYLSAKGEAF-----IG-GYIAQQAQKQKGLMTIGDVTYTKTGAVYIDSGKGGFT 240
QY 220 YIGDVRKAGEYGPPLIAGSKGSGSPMFTYDAEKQ-----KQ-----LINGILREGNPPG 271

```

Db 241 YLGGDERGGRNGIOLAKTSLKSGSTIVSGKEKGRATWMDIALIDGINAOGDZA 300
 Qy 272 KENGFOUAKSFEDEFERDLHTSLYTRAGNGVTTISGNDGSGSTQSGIPSEIKITL 331
 Db 301 KTGGF-----VETS-----GHYLSIGND--AAVEAKEMWLLDPDVTI 335
 Qy 332 ANMSLPLKEDKRVNPRVDGN-----IYSPRLNGETLTFMOKGSLIFA 378
 Db 336 SNGN---DDOSQKDRDSDPNKILADNKHVYNNKTLSTALAKGIVNISAKKCVNTAD 392
 Qy 379 SDINOGA-----GGLYFEGNFTVSPNSNOTWQ--GAGIVHNSNV----- 417
 Db 393 INVHNGTLTHSEOGGVEINGDITSPQNKTLTLAGSWVYVHAKNTITMGFLNTAGGSV 452
 Qy 418 -----TWKNGVEHRLSKIGKGLTHVQAKENGK 447
 Db 453 APERAGDGKRAAPAKIYAQVITAGSGODFRNNVS---INGGRGLKFTKMG--NGG 508
 Qy 448 SISVG-----DGKVLIEGQAD-----QNKQAFSEIGVSGRTVQ 484
 Db 509 NPSAKPDVLANISGNISINHTANNQUSYFHROGYTNWLTOLVSDSFSILTSIXDAIK 568
 Qy 485 LN-----DDKQ-----FDTKFEFGGRLDNGSLTFKRIQ-----TD 521
 Db 563 VGGYDAKDKKNTGGIGFTRDTIFNVKQAGAVDIS-YTLPIPVKNSIAAVNPDGNITV 627
 Qy 522 EGAMIVN-----HNTTOANVTITG----- 541
 Db 628 KGGGVNKLKFNALSNNKCEGWNISRPINTGSSQJLNTGSPSTLEWVANDLITNAT 687
 Qy 542 -----NESIVLPCNG-----NINKLDYKKEIAV---NOMFEE 570
 Db 688 NSPVSKIEIGTCHDITGLKYNANVTIKGSNVTLLSGNAKTKPKNVTYKCANITLAS 747
 Qy 571 TDKNHNGRLNLYKPTEDRTLLSGTNLKGDIT-----QTKGLPFSGRTPH 621
 Db 748 ANFGHKALITVAGNITQOKLAVGTDTIDVSDFTVGDAPFNNTNMLNTGFTTN 807
 Qy 622 AVN-----EAKRSEWBGIPQGEIWDHDMINEFTFAENF 657
 Db 808 GSIIDVKKGAUKGNITNEGSLNTTTHAN--TNOKTILITGNITKKGLINRDKXNA 864
 Qy 658 QIKGSAVAVSNVSSIEGNTVS-----NNANATPPVANOQNTICIR- 700
 Db 865 ELQIGG---K-SQKRGMLTISSDKVNITKQITTAGVNGENSDSGTENNANMLITKXT 919
 Qy 701 -----SDWTGLTTOQV---DLTDFKV----- 719
 Db 920 LELTNLNLISGFHKAELTAKONSDLIIGKASSDSGNAAGQVIEDKXKDSKISAGNHTV 979
 Qy 726 INSIPKQINGSLNTDANATANVAGL-----AKLMGNVTLTNHSOFTLSNNAQTGNIR 773
 Db 980 LNSEVEVS-NKSNNAAGDSNGMGLTISAKDVAVANNIT--SHRTINISAT-TGNVT 1033
 Qy 774 LSDNST-----ATVDNANINGVHLTDSAQSLKNSHPSHQLOGDKGT--VYLEN 822
 Db 1034 TKEGTTINATGGVEVETATGDIKIGIE-SKSGGVTLTHTGDTTAVANISGIVSVTANS 1092
 Qy 823 ATWMPSDTTLQNLINNSTITLINSAYSASNNTPRRSELETTTPSAEHPNTLYWNG 882
 Db 1093 GTLTTKADSTIKG--TGSVTTLSQSDIGGITSKIVSYV-TALT-----DGLTYNG 1140
 Qy 883 --KLSG--QGTPECTSLRGVSKDKLSNABEDYILSV--ANVGE-----PEYIEQ 930
 Db 1141 GAKINATEGTAITLASSGKLTTEASSITSAKQVUDLSARDNGISGINAANVTIATGT 1200
 Qy 931 LTVESKDNQPLSDKLEFLENHVDAGALRYKL VANDGEFRLNPIKEGELHNDLVAB 990
 Db 1201 LTTYKSSINANSGLTIVINADALDGT----- 1229
 Qy 991 QABRTLEAKOVEETAKTGTGEPKVRSPRAAPAFDTLLPDOSLIMALE--AKQAEITAE 1048
 Db 1230 SGDRVT--VNAVNASGSG---SVTAVTSSVNIIGDLSITNGINLIISKNGKNTIVL 1286

Qy 1049 ORSKAKTKVRSKRAVSPDPLDOSTLFALEAEVITADPOQSEKOR-LAGEAEKORQK 1107
 Db 1281 KGAELDVAYICP-----GVASAEVEIAKRALEKQDLSPEEETLAKG 1325
 Qy 1108 DILISY--SNSAL-----SELSATVNSLVSODELRELFVDOASAVWNTIAOD 1154
 Db 1326 VSAVAFVEPNNALIVNTVQNEFTTRPSQVITI-SEGACGSSGDAVCTVAD 1378
 RESULT 13
 PCT-US04-07001-672
 / Sequence 672, Application PC/TUS0407001
 / GENERAL INFORMATION:
 / APPLICANT: BAKALETZ et al.
 / TITLE OF INVENTION: GENES OF AN OTTITIS MEDIA ISOLATE OF NONTYPABLE H. INFLUENZAE
 / FILE REFERENCE: 28335/38815A
 / CURRENT APPLICATION NUMBER: PCT/US04/07001
 / CURRENT FILING DATE: 2004-03-05
 / PRIOR APPLICATION NUMBER: US 60/453,114
 / PRIOR FILING DATE: 2003-03-06
 / NUMBER OF SEQ ID NOS: 771
 / SOFTWARE: Patencin version 3.2
 / SEQ ID NO 672
 / LENGTH: 1492
 / TYPE: PRT
 / ORGANISM: H. influenzae
 PCT-US04-07001-672
 Query Match 3.8%; Score 275; DB 1; Length 1492;
 Best Local Similarity 19.1%; Pred. No. 7.5e-07;
 Matches 274; Conservative 192; Mismatches 484; Indels 484; Gaps 62;
 Qy 42 ENKKEFTVAGQNIKYNNQGGQVSTSKTAPMLDSVSRKGVALVEVQIYVAENVG 101
 Db 219 KNEGVISVNGSISL--AGQKITISDIINPTIYSI-----AAPEKAI-----NLG 264
 Qy 102 YTDVFGAEGNPDQHRFTYKIYKRNRYKKONLHPYEDDYHNPRLHFTVTEAPIDMTSN 161
 Db 265 ----DIFAKGNIN-----VRAANINOG-----KLADSVSDSGSN 298
 Qy 162 MNGSTSDRTKTPERVRIKSGRGFWENDQKDGQ--VAGAVHYLTAGNTHORGANGVYS 219
 Db 239 IVLSAKEGAE-----IG--GVISAQNOQAKGKRLMTGDRVTLTKGAVVIDSGEGEET 351
 Qy 220 YLGGDYRKAEGYGLPLAGSGKDSGPMETIYAENQ-----KX-----LINGILREKPPRG 271
 Db 352 YLGGDERGGRNGIOLAKTSLKSGSTIVSGKEKGRATWMDIALIDGINAOGKDIA 411
 Qy 272 KENGFOUAKSFEDEFERDLHTSLYTRAGNGVTTISGNDGSGSTQSGIPSEIKITL 331
 Db 412 KTGGF-----VETS-----GHYLSIGND--AAVEAKEMWLLDPDVTI 446
 Qy 332 ANMSLPLKEDKRVNPRVDGN-----IYSPRLNGETLTFMOKGSLIFA 378
 Db 447 SNGN---DDOSQKDRDSDPNKILADNKHVYNNKTLSTALAKGIVNISAKKCVNTAD 503
 Qy 379 SDINOGA-----GGLYFEGNFTVSPNSNOTWQ--GAGIVHNSNV----- 417
 Db 504 INVHNGTLTHSEOGGVEINGDITSPQNKTLTLAGSWVYVHAKNTITMGFLNTAGGSV 563
 Qy 418 -----TWKNGVEHRLSKIGKGLTHVQAKENGK 447
 Db 564 APERAGDGKRAAPAKIYAQVITAGSGODFRNNVS---INGGRGLKFTKMG--NGG 619
 Qy 448 SISVG-----DGKVLIEGQAD-----QNKQAFSEIGVSGRTVQ 484
 Db 620 NPSAKPDVLANISGNISINHTANNQUSYFHROGYTNWLTOLVSDSFSILTSIXDAIK 679
 Qy 485 LN-----DDKQ-----FDTKFEFGGRLDNGSLTFKRIQ-----TD 521
 Db 680 VGGYDAKDKKNTGGIGFTRDTIFNVKQAGAVDIS-YTLPIPVKNSIAAVNPDGNITV 738

```

QY 522 EGAMLVN-----HNTQAAVNTTG-----541
DB 739 KGGGVNLKFNALSNRYKTCGVNISSRFIVFGBOLNFGSMPTTILEVANDLIINAT 798
QY 542 -----NESIVLPGN-----NINKLYRKEIAY-----NGWGE 570
DB 799 NQSVKXIEBGTDLHDLTGAKVNAVITIKGAVTIGASAKATKEDKAVVYKAGALLIAS 858
QY 571 TQKAKANGRLNIYKTTEDRTLLISGTVLKQDIT-----QTKGLPFSSGRPPH 621
DB 859 ANFGNHKGLTVAGNINIQGLVATGDTIDVSDGFVGNDAFNGNTNNMLNTGNFTNN 918
QY 622 AVN-----HAKRSEEGEGPQGLVWCHDKIRTKAENF 657
DB 919 GTSITDYKGAAGKAGNTNNEGSLMTTHAN---TWOKTITIGNTNKGGLTRONKNA 975
QY 658 QTKGSAVSRNVSSLEGNTVS-----NNANATFGVNPQNTICTR- 700
DB 976 EIOIGG-----NISQKGNLTISSDKNITKQITIKAGVNGNSDSTENMANLTIKTK 1030
QY 701 -----SWTGLTTCQV---DITDTKV-----719
DB 1031 JELTNMNIISGFKAETITAKONSDLIIGKASSDSGAGAQVIFDKVQSKISAGNHYT 1090
QY 720 INSIPKTCJNGSINLTDNATANYKL-----AKNGVTLTNHSOFTLSNNAQTGIR 773
DB 1091 LMSVEVTS-NQNSNAGDSNGNAGLTIKADYAVANNIT---SHKTNISAT-TGAVT 1144
QY 774 JSDNST-----ATVDNANLNGVHLTDSAQFSLNKHSHQIGSKGT---VTLEN 822
DB 1145 TWEGTINATTGVEVYAKTGIDKIGIE-SKSGVYVLTATGDTLAVNGISGNTVYANS 1203
QY 823 ATWMPSDTTLQKTLNNSITTLNLSASASNNTPRRSLTETPTTSAHEHFTLLYNG 882
DB 1204 GLTLYKADSTIKG---TGSVTTLSQSGDIGTISGKTYSV-TATLT-----DSTLYKG 1251
QY 883 --KLISG--QGTFOETSSLFQYKSDKLKLNDSAGDYILSVR--NNGKE-----PETHLE 932
DB 1252 GAKINATEGTAITLTSASGKLTTEASSSTISAKQVDLSAKGIGSINAAVNTLTNGTT 1311
QY 931 LTVESKDNOPJSDKLFKTLNHDVDAKALRYKLVNDGEPRLLNPKQELHNDLVRAE 990
DB 1312 LTVYKSSINANSGLTVINAEAKLDGTA-----1340
QY 991 QAEPTLEAKQVECTATQCGPRKRSRBAKAPFTLIPDLSLMALE--AKQAEITRAET 1048
DB 1341 SGRDTV---VANATNASGG---SVTAVTSSVNTIGDLSITNGMLTISKNGNTVVL 1391
QY 1049 QKSAKTKKVRSKRAVSPDLIDQSLFLAALFVYIDAPQOSEKDR-LAQEAKQKQK 1107
DB 1392 KGAELIDVYIQP-----GASAEVIEKRLKLEKVDLSDERETLAKLG 1436
QY 1138 DLISRY--SNSAL-----SELATVNSMLSVODELRLFLVDAQASVMTNINAD 1154
DB 1437 VSAVEFEYEPNNAITVATQNEFTTSPSSQYTI-SEKACFSSGDAAVLTNVAD 1489

```

RESULT 14

```

; Sequence 41, Application US/09806709
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YANG, Yan-Ping
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; FILE REFERENCE: 1038-1138 MIS
; CURRENT APPLICATION NUMBER: US/09/806,709
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: PCT/CA99/00938
; PRIOR FILING DATE: 1999-10-07
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07

```

```

; PRIOR APPLICATION NUMBER: 09/206,942
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 1005
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
; US-09-806-709-41

```

Query Match 3.7%; Score 269; DB 5; Length 1005;

Best Local Similarity 20.5%; Pred. No. 9.5e-07; Index 364; Gaps 54;

Matches 224; Conservative 144; Mismatches 363; Indels 364; Gaps 54;

```

QY 313 GGGSIYQKSGIPSEIKITLLNMSLPIKEDKVRHPRDGFNITYSPRLNNGETLFMDQK 372
DB 19 GRGDTNKAYSAGLTSINATLSTALQGIIEV-----NISAK--NYVAVNDVYVN 68
QY 373 GSILFPASDLNAGAGLYFEGNFTVSPNSQWQAG--IHYV-----412
DB 69 GLVHLSQEN---GVKINGNITSTONGNLTITKGVKAVDVHKKITLGMGFINITSNNIT 124
QY 413 ---ENSTVWKNVGEHNR-----LSKIGKGLTVQAKEN---KGSISVGG 454
DB 125 PEKGDVLTITKQGNINQENKOLRPSVSLMGAGLFTTANKGNHTKEDVTLN-SG 183
QY 455 KVILBQ-----QADDQ-----GNKQAFSEIGVSGGVYQVND-DKQPTD 494
DB 184 KVINQTPPHMLAPNANASADS-VANVTYTLTGNNAQFTPIKFPDSRVSVALNSGRSPAG 243
QY 495 KEFFGGRGRDLNHSITLTKRLQNTD-EGAVTYNHTVQ-----AAVTTNGES- 544
DB 244 KFY---GK---NNMKENIGDANAYEFLKSNDRITSNNRPDIQFLSNISATNGTV 294
QY 545 -----IVPNGNI-----NKLDRKEIAYNGWGE-----570
DB 295 SEPIHANLSASTELNMSLINSNVNSHSHRGNAPEIKQDLINAGSFNLKQ 354
QY 571 TDKNKNGSLNLVYKPTTEDRTL---JSG---GTNLKQDIT-QTKGLPFSSGRPPH 621
DB 355 KQFQDSYKKAIF--STHNLITLGSVNTLAGESSSNIKGININSKAVNTLQA---H 408
QY 622 A-VNHLNRPWEMGIFQGEITWDEHNTNRPFAENPQKGSANVSANVSISLEGWTVS 680
DB 409 AGTSHLDKX-----ERTLLKAVSVGKALNTIGSN-AHIDGSLSTA 448
QY 681 NNAAATFGVNPQNTICTRSDTGLTTCQKVDLTQKYNISIPYQ---INGSINTLD 736
DB 449 ESAKQK-QKINNNTI-----TG--TFNNMGADINIKQGVYLGQDITNNGNINITT 498
QY 737 NATANYGLATNGVNT-----LTNHSOFTLSNNAQV-IGNIRLS-----775
DB 499 NASVQKTI--INGNITNKKGLNLIKDKANAEIOCGNISQKQENLTISSDYNITKRI 556
QY 776 -----DNSTATVDNANL-----KANVHLNDSQFSLK-----802
DB 557 EIKADTDQNSDSGVASNNALTIKTKVELITLTDNLNTSGNKEIYAKONSDLIIGASSD 616
QY 803 NSHFHSHQIGGX-----GTVTLNATWTPSDTLTQW-----LITANSITLL 845
DB 617 NSW-AKQITPQVMDSKISAGNHVNTLNSKVEITSNDSGTGSGDNNITGILISADVTV 675
QY 846 NS-----AYSSASNTPRRSLETETPTTSAHERFNTLVNKGSLQGGTQFISLF 897
DB 676 NSNITSHKTWNISASBQGITTYAGTTINATGSEVYAKTGDISGTLISG-----724
QY 898 GYKSDKLSNDARGVITSVANTGKEPETLEQVLVESKQNPQSLDKLFETLNDHYDA 957
DB 725 --KIVSVATLDS-----LTVGKAKINABGTAITLTSAGS-----KLITFANSALS 769
QY 958 GAFRYKLVNDGEPRLLNPIKEOBLHNDLVRAEOAERTLEAKOVEPTAK---TQTEBPK 1013
DB 770 GA-----NGVTASQSGDI-----SGTISGKTYSVYASSGSLTVGSDAK 808

```

QY 1014 VSRRAARAP-----DTLPDSLNALAEK-Q 1041
 DB 809 INATEGALITATKGLITTVKGSNIAMEGTLVINADATLNDASGRTIVANVNASGS 868
 QY 1042 AELTASTOKSKAKT-----KYRSRAVFSDDLQSLF--ALEALAEVIDAP 1087
 DB 869 GNVTAKTSSSVITGDSLTINGLNIISKNGKVTVLKGAIDVKYIOGVASANEVIEAK 928
 QY 1088 QOSEKDR-LAOEBAKOKOKOLISRY--SNSAL-----SELSATVNSMLSVODEJDLF 1139
 DB 929 FALKVKDLSDEBETLAKLSVSAVAFIEPNNITVTQOEFTTPSSQVTL-SEKACF 987
 QY 1140 VDOASAVMTINAD 1154
 DB 988 SSGNAAVCTINAD 1002

RESULT 15
 US-09-806-709-39
 ; Sequence 39, Application US/09866709
 ; GENERAL INFORMATION:
 ; APPLICANT: LOOSMORE, Sheera K.
 ; APPLICANT: YANG, Yan-Ping
 ; APPLICANT: KLEIN, Michel H.
 ; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
 ; TITLE OF INVENTION: MOLECULAR WEIGHT PROTEINS
 ; FILE REFERENCE: 1038-1118 MTS
 ; CURRENT APPLICATION NUMBER: US/09/806,709
 ; CURRENT FILING DATE: 2001-04-04
 ; PRIOR APPLICATION NUMBER: PCT/CA99/00938
 ; PRIOR FILING DATE: 1999-10-07
 ; PRIOR APPLICATION NUMBER: 09/167,568
 ; PRIOR FILING DATE: 1998-10-07
 ; PRIOR APPLICATION NUMBER: 09/206,942
 ; PRIOR FILING DATE: 1999-12-08
 ; NUMBER OF SEQ ID NOS: 95
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 39
 ; LENGTH: 1011
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 US-09-806-709-39

Query Match 3.7%; Score 269; DB 5; Length 1011;
 Best Local Similarity 20.5%; Pred. No. 9.6e-07;
 Matches 224; Conservative 144; Mismatches 363; Indels 364; Gaps 54;

QY 313 GQGSITOKSGIPSEIKITLANKSLPEKDKVHNPYDGNIVSPRLNSETLYPMDQK 372
 DB 25 GRGDPNKVSAEGLTSTNNALSTALQKGIY-----NISATK--WTVRAADVAKN 74
 QY 373 SGLIAPDINGAGCGIYFEGNFTVSPNSQWQAG--HVS----- 412
 DB 75 GTLVASQKN-----GVKINGNITSYQKGNLTITGKVDKVKHNTLGMGLNITSNNIT 130
 QY 413 -----ENSTVTKVNGVEHNR-----LSKIGSGTLHQAQGEN-----KGSISVDDG 454
 DB 131 FEKGNLITTAQKRI--SNQENKQIRFSNVSILNGMGAGLFTTANKGNTHKFDGTLNI-SG 189
 QY 455 KVLLEQ-----QADDQ-----GNKQAFSEIGLVSGRGVQJLND--DKQPTD 494
 DB 190 KVINQITPHNIAFWNASADSVNVTTLVLGNNAQFTFIKFDVSNMSVALNNGSGSPAGV 249
 QY 495 KFYGFGRGLDINGSLTEKRIQNTD--EGAMIVNHNTO-----AANYITGNES- 544
 DB 253 KFY-----GK-----NNEKFNIGDNAVEFKLSNDNTSNKPLPIQIFLSNISATGNQTV 300
 QY 545 -----IVLPNGNI-----NKLDYREKLVNMGFG----- 570
 DB 301 SFDIHAULSARSTELMMSLINISGVFNSHVRGNNAFAFIKKDLINATSGNENIKOT 360
 QY 571 TDKRKHGRLNLIYPTTEPRTL-----LSG-----GTNLKGDIT--QTKGKLFFSGRPPLPH 621

DB 361 KQKFDNSTERNALEF--STNLTITIGSVTLIGENSSSNIKKININKANVTLOA-----H 414
 QY 622 A-VNHNKRSSEWEKIPQGEIWDHWINRTFYAEVFCIQGSAVARSWSIEGNWTVS 680
 DB 415 AGTSLDLK-----ERTLIGVSVGNLNIISN-AHIDNLSIA 454
 QY 681 NNAATFGVVPNOQNTICTRSDWGLTTCQKVDLDTQKVISPKTO-----INGSLNLT 736
 DB 455 BSARKQ--GKTNNMLNI-----TG--TFTNGTADINIKGVVVLQGDITNNGMLNITT 504
 QY 737 NNTANVGLAKANGNT-----LTHSQFTLSNNAQ--IGNRIS----- 775
 DB 505 NASTVQKTI--INGNTKKGDINIKDIXANMEIQIGENIQKGNLTSSDKINITERI 562
 QY 776 -----DNSTATVDNANL-----NGNHLTDSNQSFLK----- 802
 DB 563 EIXADTPQNSDGVASNAANLTIKRELITLTDNLNINSGFNAELITAKNSDDLIGKASD 622
 QY 803 NSHFSHOIQGDK-----GTVTLNATVWTPSDTLQN-----LTLNNTITL 845
 DB 623 NSN-AKQITFDKVKOSKISAGNHVTLNSKVEFNSDSTGKSDDNNGLTISADVTY 681
 QY 846 NS-----AVSASSNTPRRRLBTETTPTSARENTLTVNGKLSGQGTPOFTSSLF 897
 DB 682 NSNITSKTVNISASGGITTKAFTTINATTSVEVTAKTADISGRTISG----- 730
 QY 898 GKSDXKLKSNDAAGDYILSVRNTGKEPELQLTLVESKNOQPLSDKXKFTLENDHYDA 957
 DB 731 -KTVSTLATDS-----LVKGAKINIBEGATLTLSSG-----KLTEANSALIS 775
 QY 958 GALVYKLVKQDEFLNHPYEQELANDVRAEQERTLEAKOYEPTAK-----TGTGEPK 1013
 DB 776 GA-----NGVTASSQSGDI-----SGTISGKTIVSYASSGSLVVGDAK 814
 QY 1014 VSRRAARAP-----DTLPDSLNALAEK-Q 1041
 DB 815 INATEGALITATKGLITTVKGSNIAMEGTLVINADATLNDASGRTIVANVNASGS 874
 QY 1042 AELTASTOKSKAKT-----KYRSRAVFSDDLQSLF--ALEALAEVIDAP 1087
 DB 875 GNVTAKTSSSVITGDSLTINGLNIISKNGKVTVLKGAIDVKYIOGVASANEVIEAK 934
 QY 1088 QOSEKDR-LAOEBAKOKOKOLISRY--SNSAL-----SELSATVNSMLSVODEJDLF 1139
 DB 935 FALKVKDLSDEBETLAKLSVSAVAFIEPNNITVTQOEFTTPSSQVTL-SEKACF 993
 QY 1140 VDOASAVMTINAD 1154
 DB 994 SSGNAAVCTINAD 1002

Search completed: June 23, 2004, 17:37:52
 Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2004, 17:29:18 ; Search time 26 Seconds

(Without alignments)
2767,946 Million cell updates/sec

Title: US-10-645-655-2

Perfect score: 7273

Sequence: 1 MKTVPRLNPLTACISLGIV.....SGSGQLGQGVKGLGRW 1394

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7273	100.0	1394	US-08-296-791-2	Sequence 2, Appli
2	7273	100.0	1394	US-09-839-996-2	Sequence 2, Appli
3	7273	100.0	1394	US-10-080-505-2	Sequence 2, Appli
4	7273	100.0	1394	PCT-US95-10661A-2	Sequence 2, Appli
5	7246.5	99.6	1395	US-10-080-505-7	Sequence 7, Appli
6	6115	84.1	1434	US-10-080-505-9	Sequence 9, Appli
7	5827.5	80.1	1411	US-10-080-505-17	Sequence 17, Appli
8	5807	79.8	1436	US-10-080-505-13	Sequence 13, Appli
9	5303.5	72.9	1391	US-10-080-505-15	Sequence 15, Appli
10	5297.5	72.8	1391	US-10-080-505-11	Sequence 11, Appli
11	1673.5	23.0	1541	US-08-296-791-3	Sequence 3, Appli
12	1673.5	23.0	1541	US-09-839-996-3	Sequence 3, Appli
13	1673.5	23.0	1541	US-10-080-505-3	Sequence 3, Appli
14	1673.5	23.0	1541	PCT-US95-10661A-3	Sequence 3, Appli
15	1662.5	22.9	1545	US-08-296-791-4	Sequence 4, Appli
16	1662.5	22.9	1545	US-09-839-996-4	Sequence 4, Appli
17	1662.5	22.9	1545	US-10-080-505-4	Sequence 4, Appli
18	1662.5	22.9	1545	PCT-US95-10661A-4	Sequence 4, Appli
19	1605	22.1	1702	US-08-296-791-5	Sequence 5, Appli
20	1605	22.1	1702	US-09-839-996-5	Sequence 5, Appli
21	1605	22.1	1702	US-10-080-505-5	Sequence 5, Appli
22	1605	22.1	1702	PCT-US95-10661A-5	Sequence 5, Appli
23	1599.5	21.6	1507	US-08-296-791-6	Sequence 6, Appli
24	1599.5	21.6	1507	US-09-839-996-6	Sequence 6, Appli
25	1551	21.3	1848	US-10-080-505-6	Sequence 6, Appli
26	1551	21.3	1848	US-09-839-996-6	Sequence 6, Appli
27	1551	21.3	1848	PCT-US95-10661A-6	Sequence 6, Appli

28	346.5	4.8	73	4	US-10-080-505-44	Sequence 44, Appli
29	344.5	4.7	1569	4	US-09-711-164-312	Sequence 312, App
30	286.5	3.9	1536	1	US-08-038-682-2	Sequence 2, Appli
31	286.5	3.9	1536	1	US-08-302-832-2	Sequence 2, Appli
32	286.5	3.9	1536	2	US-08-530-198-2	Sequence 2, Appli
33	286.5	3.9	1536	2	US-08-469-880-2	Sequence 2, Appli
34	286.5	3.9	1536	2	US-08-728-470-2	Sequence 2, Appli
35	286.5	3.9	1536	2	US-08-617-697-2	Sequence 2, Appli
36	286.5	3.9	1536	3	US-08-719-641-2	Sequence 3, Appli
37	285.5	3.9	1004	4	US-09-206-942-57	Sequence 57, Appli
38	285.5	3.9	1004	4	US-09-206-942-55	Sequence 55, Appli
39	283.5	3.9	1228	4	US-09-206-942-37	Sequence 37, Appli
40	283.5	3.9	1228	4	US-09-206-942-34	Sequence 34, Appli
41	283.5	3.9	1536	4	US-09-206-942-67	Sequence 67, Appli
42	283.5	3.9	1477	1	US-08-038-682-4	Sequence 4, Appli
43	283	3.9	1477	1	US-08-302-832-4	Sequence 4, Appli
44	283	3.9	1477	2	US-08-530-198-4	Sequence 4, Appli
45	283	3.9	1477	2	US-08-469-880-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-296-791-2
Sequence 2, Application US/08296791

Patent No. 6245337

GENERAL INFORMATION:

APPLICANT: St. Geme III, Joseph W.

TITLE OF INVENTION: Haemophilus Adherence and Penetration

TITLE OF INVENTION: Protein

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESSES:

ADDRESSER: Flehr, Hobach, Test, Albritton & Herbert

STREET: 4 Embarradero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: United States

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/296,791

FILING DATE: 25-AUG-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Treccartin, Richard F.

REGISTRATION NUMBER: 31,801

REFERENCE/DOCKET NUMBER: A-59941/RF7/RMS

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLCULAR TYPE: protein

Query Match 100.0%; Score 7273; DB 3; Length 1394;

Best local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;

Matches 1394; Conservative 0; Indels 0; Gaps 0;
QY 1 MKTVPRLNPLTACISLGIVSQAMAGHTYFGIDYQYBFAENKGFYGAQNIKVYVKQ 60
Db 1 MKTVPRLNPLTACISLGIVSQAMAGHTYFGIDYQYBFAENKGFYGAQNIKVYVKQ 60

```

QY 61 GQVGTSTKTAPEMIDPSVSENGVAA:VENQY:VSVANVGTVDVDFGAGNNPDQHRFT 120
DB 61 GQVGTSTKTAPEMIDPSVSENGVAA:VENQY:VSVANVGTVDVDFGAGNNPDQHRFT 120
QY 121 YKIVKRNKYNKKDLHEEDDYANRPLKFKPTEAPIDMTSNNGSYSDRTKYPERRVIG 180
DB 121 YKIVKRNKYNKKDLHEEDDYANRPLKFKPTEAPIDMTSNNGSYSDRTKYPERRVIG 180
QY 181 SGRQFENDDQKDDQVAGAYHILTAGNTHORAGNGSYLAGDVRAAGVAPLTIASCK 240
DB 181 SGRQFENDDQKDDQVAGAYHILTAGNTHORAGNGSYLAGDVRAAGVAPLTIASCK 240
QY 241 GDSGSPWETIDAEKQKMLINGILREGNPFEGKNGEQVLRKSYDELFERDHTSLYTRA 300
DB 241 GDSGSPWETIDAEKQKMLINGILREGNPFEGKNGEQVLRKSYDELFERDHTSLYTRA 300
QY 301 GNGVYTTISGNDNGQSGITQKSGIPSEIKITLNMMSLPLKEDKYNHPRYDGNIVSPRLN 360
DB 301 GNGVYTTISGNDNGQSGITQKSGIPSEIKITLNMMSLPLKEDKYNHPRYDGNIVSPRLN 360
QY 361 NGETLYEMDQKQSLIFASDINOGAGLYEGNFTVSPNSNQVQAGIHVSENGSTYWK 420
DB 361 NGETLYEMDQKQSLIFASDINOGAGLYEGNFTVSPNSNQVQAGIHVSENGSTYWK 420
QY 421 YNGVEHRLSKIGKGLHVOAKGENKSGISVGDQKYLIEQDADQGNKQAFSEIGLYSGR 480
DB 421 YNGVEHRLSKIGKGLHVOAKGENKSGISVGDQKYLIEQDADQGNKQAFSEIGLYSGR 480
QY 481 GTVQANDKQFTDXYRFGGRGLDINGHSLTFKRIQNTDEGAMVNNNTQOANVTIT 540
DB 481 GTVQANDKQFTDXYRFGGRGLDINGHSLTFKRIQNTDEGAMVNNNTQOANVTIT 540
QY 541 GNESTVLPNGNNINKLDYKELAYNGNFGHTDKNGKRLNLYKPTEDRILLSGGYN 600
DB 541 GNESTVLPNGNNINKLDYKELAYNGNFGHTDKNGKRLNLYKPTEDRILLSGGYN 600
QY 601 LKGDITQTKGLFFSGRPTPHAYNHLKRWSEMEGIPOGEIYMDHDMIRTEKAEFQIX 660
DB 601 LKGDITQTKGLFFSGRPTPHAYNHLKRWSEMEGIPOGEIYMDHDMIRTEKAEFQIX 660
QY 661 GGSVAVSRNVSISZKAWYSSNNANATGCVVQONTITCRSDWGLTQCKDLDTDFVI 720
DB 661 GGSVAVSRNVSISZKAWYSSNNANATGCVVQONTITCRSDWGLTQCKDLDTDFVI 720
QY 721 NSIPTQINGINLTJDNATAYKGLAKNGNTLTJTHSQFTLSNNAQGNIRLSDNSTA 780
DB 721 NSIPTQINGINLTJDNATAYKGLAKNGNTLTJTHSQFTLSNNAQGNIRLSDNSTA 780
QY 781 TVDNANINGVHLTJDSAQFSLKNSHFSHQIQDQKGTIVYLENAWTMPSDTTLQVLTLLN 840
DB 781 TVDNANINGVHLTJDSAQFSLKNSHFSHQIQDQKGTIVYLENAWTMPSDTTLQVLTLLN 840
QY 841 STILINAYASASNTTPRRSLTETPTPSAHRNTLTWYKGLSQGQFTGTSLSFGYK 900
DB 841 STILINAYASASNTTPRRSLTETPTPSAHRNTLTWYKGLSQGQFTGTSLSFGYK 900
QY 901 SDKILSNDAESDYILSVYKNGKEPETLEQTLTVESKNOPISDKLKTLENDVHDGAL 960
DB 901 SDKILSNDAESDYILSVYKNGKEPETLEQTLTVESKNOPISDKLKTLENDVHDGAL 960
QY 961 RYKLVKNDGEFRLNPIKEQELHNDIVRAEQAEFTLEKQVBEPLACTYGEKYSRRAA 1020
DB 961 RYKLVKNDGEFRLNPIKEQELHNDIVRAEQAEFTLEKQVBEPLACTYGEKYSRRAA 1020
QY 1021 RAAPEDTLPDQSLNALAKQAEFTLEQTSKAKTKKYSRAVSDPLDQSLFALEAA 1080
DB 1021 RAAPEDTLPDQSLNALAKQAEFTLEQTSKAKTKKYSRAVSDPLDQSLFALEAA 1080
QY 1081 LEVIDAPQSEKORLAQEAERKQKQKDLISYSNSALSELSTAVMSVYODELDTLFF 1140
DB 1081 LEVIDAPQSEKORLAQEAERKQKQKDLISYSNSALSELSTAVMSVYODELDTLFF 1140
QY 1141 DQASAVNTINIAQDKGRIDSDAPRAYQOQKTNLQIGVQALANGKIGAVSHSRSDNTF 1200
DB 1141 DQASAVNTINIAQDKGRIDSDAPRAYQOQKTNLQIGVQALANGKIGAVSHSRSDNTF 1200

```

```

DB 1141 DQASAVNTINIAQDKGRIDSDAPRAYQOQKTNLQIGVQALANGKIGAVSHSRSDNTF 1200
QY 1201 DEQVKNHATLTMGSAQYQWMDLQGVNVGTGASAKAEQSKIRKXAINGVNASY 1260
DB 1201 DEQVKNHATLTMGSAQYQWMDLQGVNVGTGASAKAEQSKIRKXAINGVNASY 1260
QY 1261 QFRLGQLGQPFQVGNRYFERENTQSERVRYKTPSLAFRRINAGIRVDYFTFTDNISY 1320
DB 1261 QFRLGQLGQPFQVGNRYFERENTQSERVRYKTPSLAFRRINAGIRVDYFTFTDNISY 1320
QY 1321 KPYFFVNYDVSNANVQTYVNLTVLQOPFGRYWQKXVGLKASILHPOJSAFISQSSQJ 1380
DB 1321 KPYFFVNYDVSNANVQTYVNLTVLQOPFGRYWQKXVGLKASILHPOJSAFISQSSQJ 1380
QY 1381 GKOQNVGYLGYRW 1394
DB 1381 GKOQNVGYLGYRW 1394

RESULT 2
US-09-839-996-2
; Sequence 2, Application US/0983996
; Patent No. 6642371
; GENERAL INFORMATION:
; APPLICANT: St. Gene III, Joseph W.
; Falkow, Stanley
; TITLE OF INVENTION: Haemophilus Adherence and Penetration
; Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIC TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/839,996
; FILING DATE: 20-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/236,791
; FILING DATE: 25-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Tregeatlin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 761-1989
; TELEFAX: (415) 398-3249
; TELEEX: 910 277299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-839-996-2

Query Match 100.0%; Score 7273; DB 4; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKKTVIRNFIYKICISLGIIVSQAMAGHTYFGIDYQYRDPANNGKFTVGAQNIKYNNQ 60
DB 1 MKKTVIRNFIYKICISLGIIVSQAMAGHTYFGIDYQYRDPANNGKFTVGAQNIKYNNQ 60

```

QY 61 GQVGTSMKAPMIDFSVSRNGVAAALVENOYTVSAHNGYTDVDFGAEGNNDQHRFT 120
 DB 61 GQVGTSMKAPMIDFSVSRNGVAAALVENOYTVSAHNGYTDVDFGAEGNNDQHRFT 120
 QY 121 YKIVKNNYKKNLHPYEDDYHNPRLHKFYTAAPIDMTSNMGSTYSORTKYEPRVIG 180
 DB 121 YKIVKNNYKKNLHPYEDDYHNPRLHKFYTAAPIDMTSNMGSTYSORTKYEPRVIG 180
 QY 181 SGROFMRNDODKDDQVAGAYHYLTAGNTHNQAGNGSYLGGVRAKAGSYGRLPIAGSK 240
 DB 181 SGROFMRNDODKDDQVAGAYHYLTAGNTHNQAGNGSYLGGVRAKAGSYGRLPIAGSK 240
 QY 241 GDSGPMFIYDAEKQKMLINGILREGNPEGKNGFOLVRKSYFDEIFERDLHTSLYTRA 300
 DB 241 GDSGPMFIYDAEKQKMLINGILREGNPEGKNGFOLVRKSYFDEIFERDLHTSLYTRA 300
 QY 301 GNGVYTTISGNDGQSLITQSGIFSEIKITLANSILPKKXKXVNPRIYDGNISYPRIN 360
 DB 301 GNGVYTTISGNDGQSLITQSGIFSEIKITLANSILPKKXKXVNPRIYDGNISYPRIN 360
 QY 361 NSEITLYFMDQKQSLIFASDINQAGGLYFEGNFTYSRNSQGTWQAGIHVSENSYTWK 420
 DB 361 NSEITLYFMDQKQSLIFASDINQAGGLYFEGNFTYSRNSQGTWQAGIHVSENSYTWK 420
 QY 421 VNGVHDLRLSKIGKGLHQAQNGKSGISYDQKXVILBQADQGNKQAPSEIGLVSR 480
 DB 421 VNGVHDLRLSKIGKGLHQAQNGKSGISYDQKXVILBQADQGNKQAPSEIGLVSR 480
 QY 481 GYVQVNDQKQPDQKFEYFEGRGRDLNGHSLYFKRIQNTDGCAMIVNHTTQAAVTLT 540
 DB 481 GYVQVNDQKQPDQKFEYFEGRGRDLNGHSLYFKRIQNTDGCAMIVNHTTQAAVTLT 540
 QY 541 GNSIYVJPGNNINXLDYRKEIAYNGMPESTDYKXKNGRLNLYKTTEDRTLLSSGGIN 600
 DB 541 GNSIYVJPGNNINXLDYRKEIAYNGMPESTDYKXKNGRLNLYKTTEDRTLLSSGGIN 600
 QY 601 LKGDITQTKGKLFESGRPTPHAVNHLNKAWSMEGIPQGEIYVMDHDMINRTKAEFQIK 660
 DB 601 LKGDITQTKGKLFESGRPTPHAVNHLNKAWSMEGIPQGEIYVMDHDMINRTKAEFQIK 660
 QY 661 GGSVSVSRNYSLEGWTVSSNNANATFGVVPQOQTTCTRSMDTGLTQCKVLDITKVI 720
 DB 661 GGSVSVSRNYSLEGWTVSSNNANATFGVVPQOQTTCTRSMDTGLTQCKVLDITKVI 720
 QY 721 NSIPKTOINSIMLTNATANYKGLAKINGNTLTMSOFTLSNNATQIGNRLSDNSTA 780
 DB 721 NSIPKTOINSIMLTNATANYKGLAKINGNTLTMSOFTLSNNATQIGNRLSDNSTA 780
 QY 781 TYDNANLNGVHLTDSNOFSLKNSHFSHOIQGKGTYYLLENAITMPSDTTLQNLTLAN 840
 DB 781 TYDNANLNGVHLTDSNOFSLKNSHFSHOIQGKGTYYLLENAITMPSDTTLQNLTLAN 840
 QY 841 STITNSAYSASNNTPRRSLTETITTSNHRFNTLTNGKLSGGQCTQFTSSIFGK 900
 DB 841 STITNSAYSASNNTPRRSLTETITTSNHRFNTLTNGKLSGGQCTQFTSSIFGK 900
 QY 901 SDKLSINDBEGYILSVANTGKEPTELQTLTVESKONPISDKLKTLENDHVAGAL 960
 DB 901 SDKLSINDBEGYILSVANTGKEPTELQTLTVESKONPISDKLKTLENDHVAGAL 960
 QY 961 FYKLVXNDGEPFLNHPILKEOELNNDVRAEOEFTLEAKOVEPLATQGTGPKYRSRRA 1020
 DB 961 FYKLVXNDGEPFLNHPILKEOELNNDVRAEOEFTLEAKOVEPLATQGTGPKYRSRRA 1020
 QY 1021 RAAFPDTLPDQSLNALBAKQAEITAEQSKAKTKYRSKRAVSDPLTQOSFALBA 1080
 DB 1021 RAAFPDTLPDQSLNALBAKQAEITAEQSKAKTKYRSKRAVSDPLTQOSFALBA 1080
 QY 1081 LEVIDAQOGEKORLAQEAERKQKQKOLISYNSAISLSTVSMKSYODELRLFY 1140
 DB 1081 LEVIDAQOGEKORLAQEAERKQKQKOLISYNSAISLSTVSMKSYODELRLFY 1140

QY 1141 DQAGAWTMIADPRRYSOAFRAYOQKTNLROIGYOKALANGRIGAVFSSHSNDTF 1200
 DB 1141 DQAGAWTMIADPRRYSOAFRAYOQKTNLROIGYOKALANGRIGAVFSSHSNDTF 1200
 QY 1201 DQOVGNATLTMSSGFAQYQWMDLOFGVNVGTISASAKAEOSRKHRAINYGNASY 1260
 DB 1201 DQOVGNATLTMSSGFAQYQWMDLOFGVNVGTISASAKAEOSRKHRAINYGNASY 1260
 QY 1261 QFRLQGLIOFPGVNRYPFERENYOSREYRVTPSLANRNACIRVDYFTPTNISV 1320
 DB 1261 QFRLQGLIOFPGVNRYPFERENYOSREYRVTPSLANRNACIRVDYFTPTNISV 1320
 QY 1321 KPYFNVYVDSQANVTNLTLYQDFGRYKQVGLKABILHROISAFISKQSGSL 1380
 DB 1321 KPYFNVYVDSQANVTNLTLYQDFGRYKQVGLKABILHROISAFISKQSGSL 1380
 QY 1381 GKQVNVGVLGYRW 1394
 DB 1381 GKQVNVGVLGYRW 1394

RESULT 3
 US-10-080-505-2
 ; Sequence 2, Application US/10080505
 ; Patent No. 6676948
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RET/DCF/DHR
 ; CURRENT FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 1394
 ; TYPE: PR
 ; ORGANISM: Haemophilus influenzae
 ; US-10-080-505-2

Query Match 100.0%; Score 7273; DB 4; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKTTFRLNPLFACISLIGISQAMAGHTYFGIDYQYTRDPAENKGFYGAQNTKYNNQ 60
 DB 1 MKTTFRLNPLFACISLIGISQAMAGHTYFGIDYQYTRDPAENKGFYGAQNTKYNNQ 60
 QY 61 GQVGTSMKAPMIDFSVSRNGVAAALVENOYTVSAHNGYTDVDFGAEGNNDQHRFT 120
 DB 61 GQVGTSMKAPMIDFSVSRNGVAAALVENOYTVSAHNGYTDVDFGAEGNNDQHRFT 120
 QY 121 YKIVKNNYKKNLHPYEDDYHNPRLHKFYTAAPIDMTSNMGSTYSORTKYEPRVIG 180
 DB 121 YKIVKNNYKKNLHPYEDDYHNPRLHKFYTAAPIDMTSNMGSTYSORTKYEPRVIG 180
 QY 181 SGROFMRNDODKDDQVAGAYHYLTAGNTHNQAGNGSYLGGVRAKAGSYGRLPIAGSK 240
 DB 181 SGROFMRNDODKDDQVAGAYHYLTAGNTHNQAGNGSYLGGVRAKAGSYGRLPIAGSK 240
 QY 241 GDSGPMFIYDAEKQKMLINGILREGNPEGKNGFOLVRKSYFDEIFERDLHTSLYTRA 300
 DB 241 GDSGPMFIYDAEKQKMLINGILREGNPEGKNGFOLVRKSYFDEIFERDLHTSLYTRA 300
 QY 301 GNGVYTTISGNDGQSLITQSGIFSEIKITLANSILPKKXKXVNPRIYDGNISYPRIN 360
 DB 301 GNGVYTTISGNDGQSLITQSGIFSEIKITLANSILPKKXKXVNPRIYDGNISYPRIN 360
 QY 361 NSEITLYFMDQKQSLIFASDINQAGGLYFEGNFTYSRNSQGTWQAGIHVSENSYTWK 420

Db 361 NGRTLYEMDQKQSLIFASDINOGAGLIFEGNFTVSPNSQTVQAGIHVSNSTVTWK 420
 QY 421 VNGVEHDLSTKIGKGLTHVQAKNGKSSISVGGCKYTLLEQCDQKQKAFSTIGLVSER 480
 Db 422 VNGVEHDLSTKIGKGLTHVQAKNGKSSISVGGCKYTLLEQCDQKQKAFSTIGLVSER 480
 QY 481 GTVQNDKQKQFDKPFYFGFRGRLDLNGSLTEKRIQNTDEKAMTNNHTTQAAVTT 540
 Db 481 GTVQNDKQKQFDKPFYFGFRGRLDLNGSLTEKRIQNTDEKAMTNNHTTQAAVTT 540
 QY 541 GNESTVLPNGNLYKLDYRKELIANGVFGFTDKKNGKNGRLNIYPTEDRTLLSGGTN 600
 Db 541 GNESTVLPNGNLYKLDYRKELIANGVFGFTDKKNGKNGRLNIYPTEDRTLLSGGTN 600
 QY 601 LKGDITOTKGLFPGSGRPTPHAYNHLNKGSENEGIPCELTWMDHNTTFAKNEFCIX 660
 Db 601 LKGDITOTKGLFPGSGRPTPHAYNHLNKGSENEGIPCELTWMDHNTTFAKNEFCIX 660
 QY 661 GGSAYVSNVSSLEGWTVSNNAATGCVWPNQNTICTRSDPTGLTTCQKVDLDTYKVI 720
 Db 661 GGSAYVSNVSSLEGWTVSNNAATGCVWPNQNTICTRSDPTGLTTCQKVDLDTYKVI 720
 QY 721 NSIPTQINSINLTQNAIYAKGLAKNGNVTLLNHSQFTLSNNAIQIGNIRLSDNSTA 780
 Db 721 NSIPTQINSINLTQNAIYAKGLAKNGNVTLLNHSQFTLSNNAIQIGNIRLSDNSTA 780
 QY 781 TVDNANLNGVHLTDSAQSLKNSHESHQIQGKGTVTLENAVTMPSTLLQMLTNN 840
 Db 781 TVDNANLNGVHLTDSAQSLKNSHESHQIQGKGTVTLENAVTMPSTLLQMLTNN 840
 QY 841 STTILNSAYSSANNTPRRRLSETEPTPTSAERFNTLLVNGKLSGGTFCQTSLSFYK 900
 Db 841 STTILNSAYSSANNTPRRRLSETEPTPTSAERFNTLLVNGKLSGGTFCQTSLSFYK 900
 QY 901 SRIKCLSNDAEGDYLLSVANTGKEPTELEQTLIVSKYNOPLSKTKFTLENDHYDAGL 960
 Db 901 SRIKCLSNDAEGDYLLSVANTGKEPTELEQTLIVSKYNOPLSKTKFTLENDHYDAGL 960
 QY 961 RYKLVKNDGEERLNPKEGEIHNLDVRAQAEKTELEKQVEPARKTOGEEKKASRRRA 1020
 Db 961 RYKLVKNDGEERLNPKEGEIHNLDVRAQAEKTELEKQVEPARKTOGEEKKASRRRA 1020
 QY 1021 RAAPDPTLPQSLNNALEKQAEITAEQSKAKTKVRSKRAVSDPLDQSLPALAAA 1080
 Db 1021 RAAPDPTLPQSLNNALEKQAEITAEQSKAKTKVRSKRAVSDPLDQSLPALAAA 1080
 QY 1081 LEVIDAPOSEKORLAQEBEAKOKKODILSRYSNALSLSATVNSMLSVODELDELFLV 1140
 Db 1081 LEVIDAPOSEKORLAQEBEAKOKKODILSRYSNALSLSATVNSMLSVODELDELFLV 1140
 QY 1141 DOASAVMTIADQRRYSDAFRAVQCCQKTLAQIGVQKALANGRLGAVSHSRSDNTF 1200
 Db 1141 DOASAVMTIADQRRYSDAFRAVQCCQKTLAQIGVQKALANGRLGAVSHSRSDNTF 1200
 QY 1201 DEQVKNALTLTMMSEFAGYQWGDLOFCVNVVGIGISASMAEBSRKIRHKAINTGVNASTY 1260
 Db 1201 DEQVKNALTLTMMSEFAGYQWGDLOFCVNVVGIGISASMAEBSRKIRHKAINTGVNASTY 1260
 QY 1261 QERIGQGLQPYRGNRRFIRRENYQSEBRKYKTSIALPNYNAIGRVDTYFTPDINISY 1320
 Db 1261 QERIGQGLQPYRGNRRFIRRENYQSEBRKYKTSIALPNYNAIGRVDTYFTPDINISY 1320
 QY 1321 KPYFVNVYDVSNANAVQTTNLTVLOQPFGRWKQKVGALQAEILHFIQISAFISQSSQL 1380
 Db 1321 KPYFVNVYDVSNANAVQTTNLTVLOQPFGRWKQKVGALQAEILHFIQISAFISQSSQL 1380
 QY 1381 GKQONVGVKLGFRM 1394
 Db 1381 GKQONVGVKLGFRM 1394

RESULT 4
 PCT-US95-10661A-2

; Sequence 2, Application PC/TUS9510661A
 ; GENERAL INFORMATION:
 ; APPLICANT: Washington University, et al.
 ; TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Piehr, Hohbach, Teet, Albritton & Herbert
 ; STREET: 4 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States
 ; ZIP: 94111-4187
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/10661A
 ; FILING DATE: 16-AUG-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/296,791
 ; FILING DATE: 25-AUG-1994
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Treacartin, Richard F.
 ; REGISTRATION NUMBER: 31,801
 ; REFERENCE/DOCKET NUMBER: PP-59941/RFT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 781-1989
 ; TELEFAX: (415) 396-3249
 ; TELEX: 9-0 277299
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1394 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US95-10661A-2
 Query Match 100.0%; Score 7273; DB 5; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKTVERLNFPLACISLIGIVSQAMAGHTYFGIDYQYRRFAENKQFTVGAQNTKYNKQ 60
 Db 1 MKTVERLNFPLACISLIGIVSQAMAGHTYFGIDYQYRRFAENKQFTVGAQNTKYNKQ 60
 QY 61 GQVGSMTKAPMTDFSVSRNGVAALVENQYIVSAHVAGYTDVDFGAEENPDQHRET 120
 Db 61 GQVGSMTKAPMTDFSVSRNGVAALVENQYIVSAHVAGYTDVDFGAEENPDQHRET 120
 QY 121 YKIVKNNYKQNDLHPYEDDYNRPILHGFVTEAPIDMTSNNGSTYSDRTKXPERVRIG 180
 Db 121 YKIVKNNYKQNDLHPYEDDYNRPILHGFVTEAPIDMTSNNGSTYSDRTKXPERVRIG 180
 QY 181 SGRQWRNDQKGDQVAGAYHYLTAGNTHNQRGANGSYLGGDVRAKAGGYPLTAGSK 240
 Db 181 SGRQWRNDQKGDQVAGAYHYLTAGNTHNQRGANGSYLGGDVRAKAGGYPLTAGSK 240
 QY 241 GSGSGPMFTYDAEKQKMLNGILIRGPNFBEKNGQVYRSYDETFEEDLTSLYTRA 300
 Db 241 GSGSGPMFTYDAEKQKMLNGILIRGPNFBEKNGQVYRSYDETFEEDLTSLYTRA 300
 QY 301 GNGVYTLISGNDNGQSITQSGIPSEIKITLANSPLLEKQKVNAPRYDGPMTYSPRLN 360
 Db 301 GNGVYTLISGNDNGQSITQSGIPSEIKITLANSPLLEKQKVNAPRYDGPMTYSPRLN 360
 QY 361 NGRTLYEMDQKQSLIFASDINOGAGLIFEGNFTVSPNSQTVQAGIHVSNSTVTWK 420
 Db 361 NGRTLYEMDQKQSLIFASDINOGAGLIFEGNFTVSPNSQTVQAGIHVSNSTVTWK 420

QY 421 VNGVEHRLSKIGKTLHVQAKGKNGKSI SVGDGVILLEQADDOGNKQAFSEIGLVSGR 480
 DB 421 VNGVEHRLSKIGKTLHVQAKGKNGKSI SVGDGVILLEQADDOGNKQAFSEIGLVSGR 480
 QY 481 GTVQNDKQFDTDFYFGFRGRLDNGHSITFRTQIONTBEGAMVNHNTTOANVTIT 540
 DB 481 GTVQNDKQFDTDFYFGFRGRLDNGHSITFRTQIONTBEGAMVNHNTTOANVTIT 540
 QY 541 GNESTIVLPNGNNINKLDRKEIAYNMGPELTDKNHNRNLTYKPTTEDRTLLSGTN 600
 DB 541 GNESTIVLPNGNNINKLDRKEIAYNMGPELTDKNHNRNLTYKPTTEDRTLLSGTN 600
 QY 601 LKGDITQTKGKLFPSGRPTPHAYNHLNKRSEMEGIPQGEIYVMDHWINRTKAEKQIK 660
 DB 601 LKGDITQTKGKLFPSGRPTPHAYNHLNKRSEMEGIPQGEIYVMDHWINRTKAEKQIK 660
 QY 661 GGSAYVSRVSSIEGNMTVSNNAATFGVVPNOQNTICTSDMTGLTTCQKVDLDTKVI 720
 DB 661 GGSAYVSRVSSIEGNMTVSNNAATFGVVPNOQNTICTSDMTGLTTCQKVDLDTKVI 720
 QY 721 NSIPKTOINGSINTLDNATANYGLAKNGVNTLTNHSOFTLSNNAITQIGNRISDNGTA 780
 DB 721 NSIPKTOINGSINTLDNATANYGLAKNGVNTLTNHSOFTLSNNAITQIGNRISDNGTA 780
 QY 781 TVDNAMNNGNVELTDSAQFSLKNSHFSHQIQDGKGTVILENATWTMSDPTLLQNLTKN 840
 DB 781 TVDNAMNNGNVELTDSAQFSLKNSHFSHQIQDGKGTVILENATWTMSDPTLLQNLTKN 840
 QY 841 SITLINSAYSASNNTPRRSLTETPTPSAEHRNTLYNGKLSGQGTFFPSLFGYK 900
 DB 841 SITLINSAYSASNNTPRRSLTETPTPSAEHRNTLYNGKLSGQGTFFPSLFGYK 900
 QY 901 SDKLKLNDAEGDYLSVNTGKEPTELEOGLIVSEKONQPSDKLKTLENDHVNAGL 960
 DB 901 SDKLKLNDAEGDYLSVNTGKEPTELEOGLIVSEKONQPSDKLKTLENDHVNAGL 960
 QY 961 RYKLVKNDGEFRLLNPIKEQELANDLVRAEQERTLEAKQVEPTAKTQSEPKVRRAA 1020
 DB 961 RYKLVKNDGEFRLLNPIKEQELANDLVRAEQERTLEAKQVEPTAKTQSEPKVRRAA 1020
 QY 1021 RAAPFDTLPDQSLALAEAKQALAEFTQKSKATKTVSKRAVPSDPLDQSLFALAEA 1080
 DB 1021 RAAPFDTLPDQSLALAEAKQALAEFTQKSKATKTVSKRAVPSDPLDQSLFALAEA 1080
 QY 1081 LEVYIDAPQOOSKORLAQEBEAKOROKOLISRYNSALSLSATNWSLSTVOBELRLV 1140
 DB 1081 LEVYIDAPQOOSKORLAQEBEAKOROKOLISRYNSALSLSATNWSLSTVOBELRLV 1140
 QY 1141 DOAQSAYVTNIAADKRRYDSDAFRAYOQOKTNLQIGVQKALANGRIKAVFSHSRSDNTF 1200
 DB 1141 DOAQSAYVTNIAADKRRYDSDAFRAYOQOKTNLQIGVQKALANGRIKAVFSHSRSDNTF 1200
 QY 1201 DEGVKNHATLTMGSGFYQYQWMDLQFVNVVGTGIGSASRAEQSKTHRAKINYNVASY 1260
 DB 1201 DEGVKNHATLTMGSGFYQYQWMDLQFVNVVGTGIGSASRAEQSKTHRAKINYNVASY 1260
 QY 1261 QFTLQGLGLOPYFGNNTFLEKXNOSEEVKTPSLAFNRNAGIRYDTEPTNTINISV 1320
 DB 1261 QFTLQGLGLOPYFGNNTFLEKXNOSEEVKTPSLAFNRNAGIRYDTEPTNTINISV 1320
 QY 1321 KPYFZYNTVDVSNANVTITVNLTVLQOPFGYWKQKAVGLKAEILHFOISAFSKSGSGL 1380
 DB 1321 KPYFZYNTVDVSNANVTITVNLTVLQOPFGYWKQKAVGLKAEILHFOISAFSKSGSGL 1380
 QY 1381 GRCQNVGVKLGFRM 1394
 DB 1381 GRCQNVGVKLGFRM 1394

GENERAL INFORMATION:
 APPLICANT: St. Gene, Joseph W.
 TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 FILE REFERENCE: A-59941-1/RF/DCF/DRR
 CURRENT APPLICATION NUMBER: US/10/080,505
 CURRENT FILING DATE: 2002-02-22
 PRIOR APPLICATION NUMBER: US 08/296,791
 PRIOR FILING DATE: 1994-10-25
 PRIOR APPLICATION NUMBER: US 09/839,996
 PRIOR FILING DATE: 2001-04-20
 NUMBER OF SEQ ID NOS: 58
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 7
 TYPE: PRT
 LENGTH: 1395
 ORGANISM: Haemophilus influenzae
 us-10-080-505-7

Query Match 99.6%; Score 7246.5; DB 4; Length 1395;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1393; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

QY 1 MKKTVFRNLPLTACISIGIVSOAMAGTYEGIDYQYRDPFAENKGFYGAONIKYVNO 60
 DB 1 MKKTVFRNLPLTACISIGIVSOAMAGTYEGIDYQYRDPFAENKGFYGAONIKYVNO 60
 QY 61 GGLVGTSTKAPMIDPSVSRNGVAALYENQYIVSAHNVGTTVDVFGADGNPDQHRT 120
 DB 61 GGLVGTSTKAPMIDPSVSRNGVAALYENQYIVSAHNVGTTVDVFGADGNPDQHRT 120
 QY 121 YKIVRNNYKKDMLHYEDDYNPRLKFTYEAAPLDMTSMNGSTYSDDTKYPERVRIG 180
 DB 121 YKIVRNNYKKDMLHYEDDYNPRLKFTYEAAPLDMTSMNGSTYSDDTKYPERVRIG 180
 QY 181 SCROFWRNDODKQVACAYHYLZAGNTHNQAGAGNGYSYAGDVRAGEYGLPIAGSK 240
 DB 181 SCROFWRNDODKQVACAYHYLZAGNTHNQAGAGNGYSYAGDVRAGEYGLPIAGSK 240
 QY 241 GSGSEWFIYDAEKOKMLINGILRBGNPFEKENGFOIVRKSYPDEIPEBDLHTSLYTRA 300
 DB 241 GSGSEWFIYDAEKOKMLINGILRBGNPFEKENGFOIVRKSYPDEIPEBDLHTSLYTRA 300
 QY 301 GNGVYTTISGNDNGQSLITQXGISEIKITLANKSLPLKEXDKVXNRRYDGPATYSPRLN 360
 DB 301 GNGVYTTISGNDNGQSLITQXGISEIKITLANKSLPLKEXDKVXNRRYDGPATYSPRLN 360
 QY 361 NGETLYTMDQKQSLIFASDINOAGGLYFEGNFTVSPNSQVTOQAGIHYSENSTYTWK 420
 DB 361 NGETLYTMDQKQSLIFASDINOAGGLYFEGNFTVSPNSQVTOQAGIHYSENSTYTWK 420
 QY 421 VNGVEHRLSKIGKTLHVQAKGKNGKSI SVGDGVILLEQADDOGNKQAFSEIGLVSGR 480
 DB 421 VNGVEHRLSKIGKTLHVQAKGKNGKSI SVGDGVILLEQADDOGNKQAFSEIGLVSGR 480
 QY 481 GTVQNDKQFDTDFYFGFRGRLDNGHSITFRTQIONTBEGAMVNHNTTOANVTIT 540
 DB 481 GTVQNDKQFDTDFYFGFRGRLDNGHSITFRTQIONTBEGAMVNHNTTOANVTIT 540
 QY 541 GNESTIVLPNGNNINKLDRKEIAYNMGPELTDKNHNRNLTYKPTTEDRTLLSGTN 600
 DB 541 GNESTIVLPNGNNINKLDRKEIAYNMGPELTDKNHNRNLTYKPTTEDRTLLSGTN 600
 QY 601 LKGDITQTKGKLFPSGRPTPHAYNHLNKRSEMEGIPQGEIYVMDHWINRTKAEKQIK 660
 DB 601 LKGDITQTKGKLFPSGRPTPHAYNHLNKRSEMEGIPQGEIYVMDHWINRTKAEKQIK 660
 QY 661 GGSAYVSRVSSIEGNMTVSNNAATFGVVPNOQNTICTSDMTGLTTCQKVDLDTKVI 720
 DB 661 GGSAYVSRVSSIEGNMTVSNNAATFGVVPNOQNTICTSDMTGLTTCQKVDLDTKVI 720
 QY 721 NSIPKTOINGSINTLDNATANYGLAKNGVNTLTNHSOFTLSNNAITQIGNRISDNGTA 780
 DB 721 NSIPKTOINGSINTLDNATANYGLAKNGVNTLTNHSOFTLSNNAITQIGNRISDNGTA 780

```

QY 781 TYDNANLNGVHLTDSAQSLKNSHSHQIGKGTVTLLNATWPSDTLONCLAN 840
Db 781 TYDNANLNGVHLTDSAQSLKNSHSHQIGKGTVTLLNATWPSDTLONCLAN 840
QY 841 STLTNSAYKSSNNTP--RRSLTETPTTSAHSPNTLVNGLSGQGTFOPTSSJFG 898
Db 841 STLTNSAYKSSNNTP--RRSLTETPTTSAHSPNTLVNGLSGQGTFOPTSSJFG 900
QY 899 YKSDKLKLSNDAGDYTLISVRNTGKEPTELBOULIVESKDNQPSLDTLXATLNDHVAG 958
Db 901 YKSDKLKLSNDAGDYTLISVRNTGKEPTELBOULIVESKDNQPSLDTLXATLNDHVAG 960
QY 959 ALRYLVNZNDEFRLLNPIKEQELNDLVBAEQAEKTLAKQVPTAKTGTGEPVRSRR 1018
Db 961 ALRYLVNZNDEFRLLNPIKEQELNDLVBAEQAEKTLAKQVPTAKTGTGEPVRSRR 1020
QY 1019 AARAFPTLPPDGLNLALAKQAEELTAFQSKKATYKVSKEAVSPDLLDGLFALE 1078
Db 1021 AARAFPTLPPDGLNLALAKQAEELTAFQSKKATYKVSKEAVSPDLLDGLFALE 1080
QY 1079 AALEYDAPQSSKEDELAQEAERKOROKDLISRYNSALSSELATYNSMLSVQDELRL 1138
Db 1081 AALEYDAPQSSKEDELAQEAERKOROKDLISRYNSALSSELATYNSMLSVQDELRL 1140
QY 1139 EYDQASAVWNTAODKRRYSDAFRAYQOQXTNLRQIGVQKALANGRGAVFSHSDN 1198
Db 1141 EYDQASAVWNTAODKRRYSDAFRAYQOQXTNLRQIGVQKALANGRGAVFSHSDN 1199
QY 1199 TPDQVKNHATLTWMSGFAQYQKQEDLQPGNNVGTGISAKKAEQSKIRKALVYGYNA 1258
Db 1200 TPDQVKNHATLTWMSGFAQYQKQEDLQPGNNVGTGISAKKAEQSKIRKALVYGYNA 1259
QY 1259 SYQRLGQIGIOPFYGVARYPTERENYSEBRYVPTSLANVYAGIVDTPTDNI 1318
Db 1260 SYQRLGQIGIOPFYGVARYPTERENYSEBRYVPTSLANVYAGIVDTPTDNI 1319
QY 1319 SVKZYFFVNVVSVNAVQTTVNLVLOQPGRYWQKEVGLKAEILHFOISAFISQGS 1378
Db 1320 SVKZYFFVNVVSVNAVQTTVNLVLOQPGRYWQKEVGLKAEILHFOISAFISQGS 1379
QY 1379 QLGKQGVGVKLYGYM 1394
Db 1380 QLGKQGVGVKLYGYM 1395

```

RESULT 6

```

US-10-080-505-9
: Sequence 9, Application US/10080505
: Patent No. 6576948
: GENERAL INFORMATION:
: APPLICANT: St. Geme, Joseph W.
: TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
: FILE REFERENCE: A-59941-1/RT/PCF/DHR
: CURRENT APPLICATION NUMBER: US/10/080,505
: CURRENT FILING DATE: 2002-02-22
: PRIOR APPLICATION NUMBER: US 08/296,791
: PRIOR FILING DATE: 1994-10-25
: PRIOR APPLICATION NUMBER: US 09/839,996
: PRIOR FILING DATE: 2001-04-20
: NUMBER OF SEQ ID NOS: 58
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 9
: LENGTH: 1434
: TYPE: PRT
: ORGANISM: Haemophilus influenzae
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (568)..(568)
: OTHER INFORMATION: The 'Xaa' at location 568 stands for Ser, Gly, Arg, or Cys.
: NAME/KEY: misc feature
: LOCATION: (1702)..(1702)
: OTHER INFORMATION: "n" at position 1702 can be any base.

```

US-10-080-505-9

Query Match 84.1%; Score 6115; DB 4; Length 1434;
 Best Local Similarity 82.9%; Pred. No. 0;
 Matches 1194; Conservative 81; Mismatches 113; Indels 52; Gaps 13;

```

QY 1 MKTVRLNLFATGICIGIVSOAMAGHYEGIDYQYRDEAKKQFTVGAQNTKVNQK 60
Db 1 MKTVRLNLFATGICIGIVSOAMAGHYEGIDYQYRDEAKKQFTVGAQNTKVNQK 60
QY 61 GOLVTSMTKAMIDFVSVRNGVAALVENVQYTVSVANTGYTDVDFAGKSNPQDRFT 120
Db 61 GOLVTSMTKAMIDFVSVRNGVAALVENVQYTVSVANTGYTDVDFAGKSNPQDRFT 120
QY 121 YKLVKNVKNKDNLPYEDDYENRPLHKRYTEAARIPMTSNMGSYSRPTKRYEPRI 180
Db 121 YKLVKNVKNKDNLPYEDDYENRPLHKRYTEAARIPMTSNMGSYSRPTKRYEPRI 180
QY 181 SGQVFRANDQKQVAGAHYITLACNTNKGAGANGSYSLGGDYRKAGEYGPLPIAGSK 240
Db 181 SGQVFRANDQKQVAGAHYITLACNTNKGAGANGSYSLGGDYRKAGEYGPLPIAGSK 240
QY 241 GDSGPMFTYDAEKQKWLINGILRBNPBEKENGPOYRKSYFDEIPEBDLHLSITTRA 300
Db 241 GDSGPMFTYDAEKQKWLINGILRBNPBEKENGPOYRKSYFDEIPEBDLHLSITTRA 300
QY 301 GNGVYTTISGNDQGSITQKSGIPSEIKITLANNSLPLEKDKVNPYDGPNIYSPRLN 360
Db 301 GNGVYTTISGNDQGSITQKSGIPSEIKITLANNSLPLEKDKVNPYDGPNIYSPRLN 360
QY 361 NSEFTLYFMDQKQSLFASLINOAGGCIYFEGATTVSPNSNOTWQAGIHVSENSYTWK 420
Db 359 NSEFTLYFMDQKQSLFASLINOAGGCIYFEGATTVSPNSNOTWQAGIHVSENSYTWK 418
QY 421 VNGVEHRLSKIGLGLHVOAKGENKGSISYQKQVILEQCDQDQGNQAAASEGLYSR 480
Db 419 VNGVEHRLSKIGLGLHVOAKGENKGSISYQKQVILEQCDQDQGNQAAASEGLYSR 478
QY 481 GTVOLNDKQFDTRKYFGRGRJDLNGHSLTFKSLQNTDGAALVNHNTTOAAVNTT 540
Db 479 GTVOLNDKQFDTRKYFGRGRJDLNGHSLTFKSLQNTDGAALVNHNTTOAAVNTT 538
QY 541 GNESTVLP--NGNNINLQDTRKELIANGMFGETTRKNGKRLMLYKPTTEDRLILSGGT 595
Db 533 GNESTVLP--NGNNINLQDTRKELIANGMFGETTRKNGKRLMLYKPTTEDRLILSGGT 598
QY 600 NLKGDITQTKGLTFPSGRPTPHAVNHLNKRSMESGIPQGLIYDMDHINRTFKASNP 659
Db 599 NLKGDITQTKGLTFPSGRPTPHAVNHLNKRSMESGIPQGLIYDMDHINRTFKASNP 658
QY 660 KQGSAYSVRSVSSIEGNTVTSNNNAATFGVVPQDNTICTRSMTGLTTCQKVDLTDKV 719
Db 659 KQGSAYSVRSVSSIEGNTVTSNNNAATFGVVPQDNTICTRSMTGLTTCQKVDLTDKV 718
QY 720 INSLPKQINGSLNLDAATANGKLAANGNTLVNHSQTTLSNNAATQIGINISNAN 778
Db 719 INSLPKQINGSLNLDAATANGKLAANGNTLVNHSQTTLSNNAATQIGINISNAN 778
QY 776 ---DNST-----ATVDNANLNGVHLTDSAQSL 800
Db 779 ARVNATLMDVNLADTSRFLSNQATQIGTSLHQQAQATVDNANLNGVHLTDSAQSL 838
QY 801 LKNSHFSHQIQGDKGTYVLENATWPSDTTLOVLTNNSTTLNLSAYSASSNNTTP--RR 859
Db 839 LKNSHFSHQIQGDKGTYVLENATWPSDTTLOVLTNNSTTLNLSAYSASSNNTTP--RR 858
QY 860 RSLTETPTTSAHSPNTLVNGLSGQGTFOPTSSJFG 919
Db 899 RSLTETPTTSAHSPNTLVNGLSGQGTFOPTSSJFG 958
QY 920 NTGKEPTELBOULIVESKDNQPSLDTLXATLNDHVAGALRYLVNZNDEFRLLNPIKE 979
Db 959 NTGKEPTELBOULIVESKDNQPSLDTLXATLNDHVAGALRYLVNZNDEFRLLNPIKE 1018

```


QY 98C QELNDLVRAEQAEKRLTAKOYEPTACTGTGEBKYSRRARAAPPTLIPDQSLNATEA 1039
 DB 1019 QELNSDLVRAEQAEKRLTAKOYEPTACTGTGEBKYSRRARAAPPTLIPDQSLNATEA 1075
 QY 1040 KQALLETOTKSAKTKKYSKRAV---FSDLLDQS-LFLEALAEVTDAPQSEKRL 1095
 DB 1076 KQAPNAKTKSKAKTKKYSKRAV---FSDLLDQS-LFLEALAEVTDAPQSEKRL 1135
 QY 1096 AQ-ESEAKORKKQKDLISRYNSALSLSLNTVMSLSYODELRLFPDQASVMTNIAOD 1154
 DB 1136 AQ-ESEAKORKKQKDLISRYNSALSLSLNTVMSLSYODELRLFPDQASVMTNIAOD 1195
 QY 1155 KRRYSDAFRAYQOQKTLNLRQIGVQKALANGRIKAVFSHRSNDTDEBYKHAATLIMS 1214
 DB 1196 KRRYSDAFRAYQOQKTLNLRQIGVQKALANGRIKAVFSHRSNDTDEBYKHAATLIMS 1254
 QY 1215 GFAQYQMGDLQSGVNVGTGISASMAEESRKIRKALINGVNASYQFRLQGLQPTLIG 1274
 DB 1255 GFAQYQMGDLQSGVNVGTGISASMAEESRKIRKALINGVNASYQFRLQGLQPTLIG 1314
 QY 1275 VNRFFERENYQSEEVKQTPSLAFKRYNAGIRVDYPTPTDNI SYKPYEFVNVDSNA 1334
 DB 1315 VNRFFERENYQSEEVKQTPSLAFKRYNAGIRVDYPTPTDNI SYKPYEFVNVDSNA 1374
 QY 1335 NVQTYNLTVLQOQPFGRYQKEVGLKAEILHFPQISAFISKSQSGQKQNVGKGYRM 1394
 DB 1375 NVQTYNLTVLQOQPFGRYQKEVGLKAEILHFPQISAFISKSQSGQKQNVGKGYRM 1434

RESULT 7

US-10-080-505-17
 ; Sequence 17, Application US/10080505
 ; Patent No. 6676348
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Geme, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/REF/DCF/DHR
 ; CURRENT APPLICATION NUMBER: US/10/080,505
 ; CURRENT FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 17
 ; LENGTH: 1411
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 ; US-10-080-505-17

Query Match 86.1%; Score 5827.5; DB 4; Length 1411;
 Best local Similarity 79.5%; Pred. No. 0;
 Matches 1133; Conservative 107; Mismatches 139; Indels 47; Gaps 12;

QY 1 MKKTVFRLNPLTACTSLGIVSOAMAGHTYEGIDYQYRRPAENKSGKTEGAKNIKKYKQ 60
 DB 1 MKKTVFRLNPLTACTSLGIVSOAMAGHTYEGIDYQYRRPAENKSGKTEGAKNIKKYKQ 60
 QY 61 GQLVSTSKAPMIDSVSVSRNGVALVENVQTVSVANHVGVTVDFGAEGNPDQHEFT 120
 DB 61 GQLVSTSKAPMIDSVSVSRNGVALVENVQTVSVANHVGVTVDFGAEGNPDQHEFT 120
 QY 121 YKIVKNNYKKDNHAYEDDYNNPRLNHFVTEAAPTLMKSNNGSTYSDRTKYPKRVIG 180
 DB 121 YKIVKNNYKKDNHAYEDDYNNPRLNHFVTEAAPTLMKSNNGSTYSDRTKYPKRVIG 180
 QY 181 SGRQFWRNDQKEDQVAGAYHYLLAGNTHNQAGANGYSYGGVRYKAGCYGGLPLAGSK 240
 DB 181 SGRQFWRNDQKEDQVAGAYHYLLAGNTHNQAGANGYSYGGVRYKAGCYGGLPLAGSK 240
 QY 241 GHSQSPWFYDAEKQKMLINGILABNGPFGKENGQFLVRKSYDEIIFENDLHSLTYTRA 300

DB 241 GHSQSPWFYDAEKQKMLINGILABNGPFGKENGQFLVRKSYDEIIFENDLHSLTYTRA 300
 QY 301 GNGVYIISGNDQGSITOKSGIPSEIKITLANNLFLKPKKYNHPYDQPNYSPRLN 360
 DB 301 GNGVYIISGNDQGSITOKSGIPSEIKITLANNLFLKPKKYNHPYDQPNYSPRLN 360
 QY 361 NGSETLYMDQKGSILPASDINGAGGLYFEGNFTVSPNSQOTWQAGIHVSENSYVTK 420
 DB 361 NGSETLYMDQKGSILPASDINGAGGLYFEGNFTVSPNSQOTWQAGIHVSENSYVTK 420
 QY 421 VNGVEHDLSTKIGGTLHYQAKENKSGISYQDQKVLLEQDQDQKQAFSEIGVSGR 480
 DB 421 VNGVEHDLSTKIGGTLHYQAKENKSGISYQDQKVLLEQDQDQKQAFSEIGVSGR 480
 QY 481 GTVQNLNDKQPTDQPFQFGRGRLDNGSHLPEKRIQNTDQGANIVNHNQQAANYTTI 540
 DB 481 GTVQNLNDKQPTDQPFQFGRGRLDNGSHLPEKRIQNTDQGANIVNHNQQAANYTTI 540
 QY 541 GNESTVLPNGANNINKLDYKKEIYVNGFGEYDQKKNKHEKRLTYPTTEDEKTLISGGIN 600
 DB 541 GNESTVLPNGANNINKLDYKKEIYVNGFGEYDQKKNKHEKRLTYPTTEDEKTLISGGIN 600
 QY 593 GSDTL---NDNTGDLTNKRDIAFNGMGQDQDQKQNTGRLANTVYPMKQKHLLSGGIN 594
 DB 593 GSDTL---NDNTGDLTNKRDIAFNGMGQDQDQKQNTGRLANTVYPMKQKHLLSGGIN 594
 QY 601 LKGDQITQKQKLPFSGRPTPAAVNTLAKGSEMEGIPQGEIYWDHMJNTPFAENFQIK 660
 DB 601 LKGDQITQKQKLPFSGRPTPAAVNTLAKGSEMEGIPQGEIYWDHMJNTPFAENFQIK 660
 QY 653 GGSIVSVSNVSIIEGMMVSNANAATGVVNOONTICRSPTWGLTQCKQVLTPTTKYI 720
 DB 653 GGSIVSVSNVSIIEGMMVSNANAATGVVNOONTICRSPTWGLTQCKQVLTPTTKYI 720
 QY 721 NSIPKTOQINSINLTNATANTYKGLAKNGVTLTNHSPQTLSSNATQGINRLSDNSTA 780
 DB 721 NSIPKTOQINSINLTNATANTYKGLAKNGVTLTNHSPQTLSSNATQGINRLSDNSTA 780
 QY 773 TYDNANLNGVNLHLDQAFSLKNSHFSHQIQGKGTVTLENAITVHPSSTTLQNTLTK 840
 DB 773 TYDNANLNGVNLHLDQAFSLKNSHFSHQIQGKGTVTLENAITVHPSSTTLQNTLTK 840
 QY 844 STILNSAYASSNNTPR-RSLSETTTPSAHRPNTLTYNGKLSQGFQGTSLSEY 899
 DB 844 STILNSAYASSNNTPR-RSLSETTTPSAHRPNTLTYNGKLSQGFQGTSLSEY 899
 QY 893 KSDIKKLSNABSDYTLAVRDTGKEPVTLLEQLTLEGDNPFLPDKKTLTKNKHVDGA 952
 DB 893 KSDIKKLSNABSDYTLAVRDTGKEPVTLLEQLTLEGDNPFLPDKKTLTKNKHVDGA 952
 QY 960 LRYKLVKNDGEPRLNPEIKROELNDLVRAEQAEKRLTAKOYEPTACTGTGEBKYSRRA 1019
 DB 960 LRYKLVKNDGEPRLNPEIKROELNDLVRAEQAEKRLTAKOYEPTACTGTGEBKYSRRA 1019
 QY 1019 ARAAFPTLIPDQSLNATEAKQAELETAETQSKA-----KTKKYSRRA---AV 1064
 DB 1019 ARAAFPTLIPDQSLNATEAKQAELETAETQSKA-----KTKKYSRRA---AV 1064
 QY 1065 FSDLLDQS---FA-----LEALAEVTDAPQSEKRLTAQEAERKQKQD 1108
 DB 1065 FSDLLDQS---FA-----LEALAEVTDAPQSEKRLTAQEAERKQKQD 1108
 QY 1109 LISRYNSALSLSLNTVMSLSYODELRLFPDQASVMTNIAODKRRYSDAFRAYQ 1168
 DB 1109 LISRYNSALSLSLNTVMSLSYODELRLFPDQASVMTNIAODKRRYSDAFRAYQ 1168
 QY 1127 LISRYNSALSLSLNTVMSLSYODELRLFPDQASVMTNIAODKRRYSDAFRAYQ 1185
 DB 1127 LISRYNSALSLSLNTVMSLSYODELRLFPDQASVMTNIAODKRRYSDAFRAYQ 1185
 QY 1169 QKTLNLRQIGVQKALANGRIKAVFSHRSNDTDEBYKHAATLIMS6FAQYQMGDLQPGV 1228
 DB 1169 QKTLNLRQIGVQKALANGRIKAVFSHRSNDTDEBYKHAATLIMS6FAQYQMGDLQPGV 1228
 QY 1229 NVGNGISASMAEESRKIRKALINGVNASYQFRLQGLQYFVNVNRYFERENYQSE 1288
 DB 1229 NVGNGISASMAEESRKIRKALINGVNASYQFRLQGLQYFVNVNRYFERENYQSE 1288
 QY 1246 NVGNGISASMAEESRKIRKALINGVNASYQFRLQGLQYFVNVNRYFERENYQSE 1305
 DB 1246 NVGNGISASMAEESRKIRKALINGVNASYQFRLQGLQYFVNVNRYFERENYQSE 1305
 QY 1289 EVRKTPSLAFKRYNAGIRVDYPTPTDNI SYKPYEFVNVDSNANVQTYNLTVLQOP 1348
 DB 1289 EVRKTPSLAFKRYNAGIRVDYPTPTDNI SYKPYEFVNVDSNANVQTYNLTVLQOP 1348

Db 1306 EVKVCPTSLAFNRNAGVRVDVTFPTENISVKEPFVNVYDVSNANCTTVRAVLQOP 1365

QY 1349 FGRYQCKEYGLKAEILHFQISAFISKSQSLQKQOYGVKLGYYM 1394

Db 1366 FGRYQCKEYGLKAEILHFQISAFISKSQSLQKQOYGVKLGYYM 1411

RESULT 8

US-10-080-505-13

/ Sequence 13, Application US/10080505

/ Patent No. 6676948

/ GENERAL INFORMATION:

/ APPLICANT: St. Gene, Joseph W.

/ TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS

/ FILE REFERENCE: A-59941-1/RT/DCF/DHR

/ CURRENT APPLICATION NUMBER: US/10/080,505

/ CURRENT FILING DATE: 2002-02-22

/ PRIOR APPLICATION NUMBER: US 08/296,791

/ PRIOR FILING DATE: 1994-10-25

/ PRIOR APPLICATION NUMBER: US 09/839,996

/ PRIOR FILING DATE: 2001-04-20

/ NUMBER OF SEQ ID NOS: 58

/ SOFTWARE: Patent version 3.1

/ SEQ ID NO 13

/ LENGTH: 1436

/ TYPE: PRT

/ ORGANISM: Haemophilus influenzae

US-10-080-505-13

Query Match 79.8%; Score 5807; DB 4; Length 1436;

Best Local Similarity 79.3%; Pred. No. 0;

Matches 1149; Conservative 78; Mismatches 154; Indels 66; Gaps 14;

QY 1 MKTVFLNLTFLACISLGIVSOAMAGHTVGIQYRDPFAKNGKFTVGAQNTVYKQ 60

Db 1 MKTVFLNLTFLACISLGIVSOAMAGHTVGIQYRDPFAKNGKFTVGAQNTVYKQ 60

QY 61 GQVGSMTKAPMIDSVYSRNGVALVENQYVSAANVETDVFAGAGNPPORHFT 120

Db 61 GQVGSMTKAPMIDSVYSRNGVALVENQYVSAANVETDVFAGAGNPPORHFT 120

QY 121 YKIKNNYKCNLHPEYEDDYNHRLHKTVAATLIMTSNNGSTYSRTYPERVYIG 180

Db 121 YKIKNNYKCNLHPEYEDDYNHRLHKTVAATLIMTSNNGSTYSRTYPERVYIG 180

QY 122 YQIVKNNYKPGKDPYHGDYHMRHKEVTLAEPAKWTQNMKGKXADLSYDPRVRIG 180

Db 122 YQIVKNNYKPGKDPYHGDYHMRHKEVTLAEPAKWTQNMKGKXADLSYDPRVRIG 180

QY 181 SGQFWRNDP-----KGDYAGAYHYLTKAGTNHORGAGVSYLGGVYKRAEYEPD 235

Db 181 SGQFWRNDP-----KGDYAGAYHYLTKAGTNHORGAGVSYLGGVYKRAEYEPD 235

QY 236 TAGSKDSSPMFYDAEKOKMLINGLBEKNPEEGKNGFQIVRKSYF-DEIFERDLHT 294

Db 236 TAGSKDSSPMFYDAEKOKMLINGLBEKNPEEGKNGFQIVRKSYF-DEIFERDLHT 294

QY 241 TGVSFGDSGPMFYDAIKQKWLINGVLQTNPFSGANGFQILRNKMFYDNVFEVDELPI 300

Db 241 TGVSFGDSGPMFYDAIKQKWLINGVLQTNPFSGANGFQILRNKMFYDNVFEVDELPI 300

QY 295 SLVFRANGVYITISNDNGOSITQ--KSGIPSEIKITLANNGLPEKOKVHNPRY-- 349

Db 295 SLVFRANGVYITISNDNGOSITQ--KSGIPSEIKITLANNGLPEKOKVHNPRY-- 349

QY 301 TTEPRNGHISFTISNNGTGVTQYNEKVSMPQKVTATQULFENALKEKCK--EPYAA 358

Db 301 TTEPRNGHISFTISNNGTGVTQYNEKVSMPQKVTATQULFENALKEKCK--EPYAA 358

QY 350 DGPNIYSPRIANGELTYEMDQKQSLIFASDINOAGGLYFEGSNFYSPNSNQWAGGI 409

Db 350 DGPNIYSPRIANGELTYEMDQKQSLIFASDINOAGGLYFEGSNFYSPNSNQWAGGI 409

QY 359 GGVNAYKPRILNKGKIIYGRDGRGCTLTITNNINQAGGLHFSNFTVSENNATWQAGAG 413

Db 359 GGVNAYKPRILNKGKIIYGRDGRGCTLTITNNINQAGGLHFSNFTVSENNATWQAGAG 413

QY 410 HYSBNSTVYKNGVYEHDLSTKIGKTLHVAQKENGKSIWGDVILAEQADDOGNKQ 469

Db 410 HYSBNSTVYKNGVYEHDLSTKIGKTLHVAQKENGKSIWGDVILAEQADDOGNKQ 469

QY 419 HVEDSTVYKNGVYEHDLSTKIGKTLHVAQKENGKSIWGDVILAEQADDOGNKQ 478

Db 419 HVEDSTVYKNGVYEHDLSTKIGKTLHVAQKENGKSIWGDVILAEQADDOGNKQ 478

QY 470 AFSEGLVSGRGTVQLNDNDQOEDTQKYGGRGRDLNLSHTTKRLQNTDQAGATVNH 529

Db 470 AFSEGLVSGRGTVQLNDNDQOEDTQKYGGRGRDLNLSHTTKRLQNTDQAGATVNH 529

QY 479 AFKEVGIVSGRATVQLNSADQVDPNNLYFGRGRDLNLSHTTKRLQNTDQAGATVNH 538

Db 479 AFKEVGIVSGRATVQLNSADQVDPNNLYFGRGRDLNLSHTTKRLQNTDQAGATVNH 538

QY 530 NTTQANVTITGNESVLP--NGNNIKLIDYKKEIAYVNGFGETDKKNGKNGVNTIKYPTT 588

Db 530 NTTQANVTITGNESVLP--NGNNIKLIDYKKEIAYVNGFGETDKKNGKNGVNTIKYPTT 588

QY 539 NTTQANVTITGNESVLP--NGNNIKLIDYKKEIAYVNGFGETDKKNGKNGVNTIKYPTT 598

Db 539 NTTQANVTITGNESVLP--NGNNIKLIDYKKEIAYVNGFGETDKKNGKNGVNTIKYPTT 598

QY 589 EDRTLLISGCTNKQDITQYKGLFESGRPTPAVNLNKRKSEMGISDQEIYWDHDI 648

Db 589 EDRTLLISGCTNKQDITQYKGLFESGRPTPAVNLNKRKSEMGISDQEIYWDHDI 648

QY 599 EDRTLLISGCTNKQDITQYKGLFESGRPTPAVNLNKRKSEMGISDQEIYWDHDI 656

Db 599 EDRTLLISGCTNKQDITQYKGLFESGRPTPAVNLNKRKSEMGISDQEIYWDHDI 656

QY 649 NRTKAPNPOIGGSANVSRVSSIEGNTVSNANNAIFGVVPOQNTICTSDWTGLTT 708

Db 649 NRTKAPNPOIGGSANVSRVSSIEGNTVSNANNAIFGVVPOQNTICTSDWTGLTT 708

QY 657 TRTKAPNPOIGGSANVSRVSSIEGNTVSNANNAIFGVVPOQNTICTSDWTGLTT 716

Db 657 TRTKAPNPOIGGSANVSRVSSIEGNTVSNANNAIFGVVPOQNTICTSDWTGLTT 716

QY 703 QOKVLDITDTEVINSIFPTQINGSINTLNATNATVNGKANGVTLTNSSQFTISNNA 768

Db 703 QOKVLDITDTEVINSIFPTQINGSINTLNATNATVNGKANGVTLTNSSQFTISNNA 768

QY 717 KCTVDLIDITKIVINSIFPTQINGSINTLNATNATVNGKANGVTLTNSSQFTISNNA 776

Db 717 KCTVDLIDITKIVINSIFPTQINGSINTLNATNATVNGKANGVTLTNSSQFTISNNA 776

QY 769 IGTIELSDNSTATVDNANLNGNHLTDSAQFSLKNSHSHQIQDGTITLNNATVMP 828

Db 769 IGTIELSDNSTATVDNANLNGNHLTDSAQFSLKNSHSHQIQDGTITLNNATVMP 828

QY 777 TGNIOJLSNANNAITVDNANLNGNHLTDSAQFSLKNSHSHQIQDGTITLNNATVMP 836

Db 777 TGNIOJLSNANNAITVDNANLNGNHLTDSAQFSLKNSHSHQIQDGTITLNNATVMP 836

QY 829 SDTLQNLNANSTTILNNSAYASASNNTR--RSLFETTTPTSAHRFVTLTVNGKLSGQ 887

Db 829 SDTLQNLNANSTTILNNSAYASASNNTR--RSLFETTTPTSAHRFVTLTVNGKLSGQ 887

QY 837 SDTLQNLNANSTTILNNSAYASASNNTR--RSLFETTTPTSAHRFVTLTVNGKLSGQ 896

Db 837 SDTLQNLNANSTTILNNSAYASASNNTR--RSLFETTTPTSAHRFVTLTVNGKLSGQ 896

QY 888 GTEQFTSSLFYKSKDKLNSDAEGDYILSVANNKEPETLAEQTLVESKONOP--SDK 947

Db 888 GTEQFTSSLFYKSKDKLNSDAEGDYILSVANNKEPETLAEQTLVESKONOP--SDK 947

QY 897 GTEQFTSSLFYKSKDKLNSDAEGDYILSVANNKEPETLAEQTLVESKONOP--SDK 956

Db 897 GTEQFTSSLFYKSKDKLNSDAEGDYILSVANNKEPETLAEQTLVESKONOP--SDK 956

QY 948 FTLENDHNDAGALKYKYKNDGFEFLNPIYQELANDIVRABQARTLEAKQVEPTAKT 1007

Db 948 FTLENDHNDAGALKYKYKNDGFEFLNPIYQELANDIVRABQARTLEAKQVEPTAKT 1007

QY 957 FTLENDHNDAGALKYKYKNDGFEFLNPIYQELANDIVRABQARTLEAKQVEPTAKT 1016

Db 957 FTLENDHNDAGALKYKYKNDGFEFLNPIYQELANDIVRABQARTLEAKQVEPTAKT 1016

QY 1008 QTEGPKVSRBARAPDPTLPPOSILNNAEAKNELTETQSKAKTKKVRKRAVFS 1067

Db 1008 QTEGPKVSRBARAPDPTLPPOSILNNAEAKNELTETQSKAKTKKVRKRAVFS 1067

QY 1017 QTNANAVSK--RAVSDTLTPDQSDVLDQAEQVEPTAKQKNNNA--KKRSGRAVFS 1071

Db 1017 QTNANAVSK--RAVSDTLTPDQSDVLDQAEQVEPTAKQKNNNA--KKRSGRAVFS 1071

QY 1068 PLIDQS-----LEALAEALVID 1085

Db 1068 PLIDQS-----LEALAEALVID 1085

QY 1072 TLPDQSDVLDQAEQVEPTAKQKAKKAKKAKKAAEESDPTLPDLSLAKYLEKLVIN 1131

Db 1072 TLPDQSDVLDQAEQVEPTAKQKAKKAKKAKKAAEESDPTLPDLSLAKYLEKLVIN 1131

QY 1086 APQOSKRLAQBEBEAKQKQDLISYSNALSSELSATVNSLVSQDELDTLPVDAQS 1145

Db 1086 APQOSKRLAQBEBEAKQKQDLISYSNALSSELSATVNSLVSQDELDTLPVDAQS 1145

QY 1132 AQOQVKE--PDQEKQRKQDLISYSNALSSELSATVNSLVSQDELDTLPVDAQS 1188

Db 1132 AQOQVKE--PDQEKQRKQDLISYSNALSSELSATVNSLVSQDELDTLPVDAQS 1188

QY 1146 AVWTNLADKRYSPDAPRAYOQKTNRLQIGVQYALANGRGVFSHSSNDTPPEQYK 1205

Db 1146 AVWTNLADKRYSPDAPRAYOQKTNRLQIGVQYALANGRGVFSHSSNDTPPEQYK 1205

QY 1189 AVWTNLADKRYSPDAPRAYOQKTNRLQIGVQYALANGRGVFSHSSNDTPPEQYK 1247

Db 1189 AVWTNLADKRYSPDAPRAYOQKTNRLQIGVQYALANGRGVFSHSSNDTPPEQYK 1247

QY 1206 NHAATLTMGSPAYQWGDQFCGVNVTGISASVMAEESQRKHKRALYGVNASYQFRLG 1265

Db 1206 NHAATLTMGSPAYQWGDQFCGVNVTGISASVMAEESQRKHKRALYGVNASYQFRLG 1265

QY 1248 NHAATLTMGSPAYQWGDQFCGVNVTGISASVMAEESQRKHKRALYGVNASYQFRLG 1307

Db 1248 NHAATLTMGSPAYQWGDQFCGVNVTGISASVMAEESQRKHKRALYGVNASYQFRLG 1307

QY 1266 QLGIGPYGVNRYFIEREYVQSEEVRYKTPSLAPRYNAGLRVDYTPPTDNI SVAPYF 1335

Db 1266 QLGIGPYGVNRYFIEREYVQSEEVRYKTPSLAPRYNAGLRVDYTPPTDNI SVAPYF 1335

QY 1308 QLGIGPYGVNRYFIEREYVQSEEVRYKTPSLAPRYNAGLRVDYTPPTDNI SVAPYF 1367

Db 1308 QLGIGPYGVNRYFIEREYVQSEEVRYKTPSLAPRYNAGLRVDYTPPTDNI SVAPYF 1367

QY 1326 VNYVDSNANVQTTVALTVLQOPFGRYQKVEGLKAEILHFQISAFISKSQSLQKQON 1385

Db 1326 VNYVDSNANVQTTVALTVLQOPFGRYQKVEGLKAEILHFQISAFISKSQSLQKQON 1385

QY 1368 VNYVDSNANVQTTVALTVLQOPFGRYQKVEGLKAEILHFQISAFISKSQSLQKQON 1427

Db 1368 VNYVDSNANVQTTVALTVLQOPFGRYQKVEGLKAEILHFQISAFISKSQSLQKQON 1427

RESULT 9

US-10-080-505-15

/ Sequence 15, Application US/10080505

/ Patent No. 6676948

/ GENERAL INFORMATION:

/ APPLICANT: St. Gene, Joseph W.

/ TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS

/ FILE REFERENCE: A-59941-2/RT/DCF/DHR

/ CURRENT APPLICATION NUMBER: US/10/080,505

/ PRIOR APPLICATION NUMBER: US 08/296,791

/ PRIOR FILING DATE: 1994-10-25

; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patent In version 3.1
 ; SEQ ID NO 15
 ; LENGTH: 1391
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 us-10-080-505-11

Query Match 72.8%; Score 5303.5; DB 4; Length 1391;
 Best Local Similarity 74.8%; Pred. No. 0;
 Matches 1053; Conservative 109; Mismatches 215; Indels 31; Gaps 15;

QY 1 MKTIVRLNLTACISLIGIVSQAAMAGHTYEGIDYQYRDBAENKGFYGAQNIKYNNKQ 60
 DB 1 MKTIVRLNLTACISLIGIVSQAAMAGHTYEGIDYQYRDBAENKGFYGAQNIKYNNK 60
 QY 61 GQLVGSKTKAPMDIFSVSNGVAALVENQYIVSAHVNGYTDVDFGAEGNPDQHRFT 120
 DB 61 GEMIGTMKGVPMPDLSSMVRGYSSTLSEQHLLSVAHNGYDVDFGEGENPDQHRFX 120
 QY 121 YKIVKNNYKKNLHFEEDYHNPRLHFKFTBAPPIVTSNNXGSSYSDRTKYPERVYIG 180
 DB 121 YKIVKNNYKKNLHFEEDYHNPRLHFKFTBAPPIVTSNNXGSSYSDRTKYPERVYIG 178
 QY 181 SGHOFNRNDQK--GDQVAGAYYLTAGTINHQAGNGSYLGGDVKAGEVPLPIAG 238
 DB 179 SGHOFNRNDQK--GDQVAGAYYLTAGTINHQAGNGSYLGGDVKAGEVPLPIAG 237
 QY 239 SKGDSGSEPFYDAEKOKMLINGILREGNPEEGKENGFOIVKRSYFDEIFERDLFTSLYT 298
 DB 238 SKGDSGSEPFYDAEKOKMLINGILREGNPEEGKENGFOIVKRSYFDEIFERDLFTSLYT 297
 QY 299 BAANGYTTISANDG-----QSGITQKSGIPSEIKITLANWSL-PLKEDKVNHPRTDGP 352
 DB 298 BAANGYTTISANDG-----QSGITQKSGIPSEIKITLANWSL-PLKEDKVNHPRTDGP 352
 QY 353 NIYSPLNNGETLYLFMOQKQSLIFASDINQAGCELYFGCNTTYSRNSQ--TNGQGHYV 411
 DB 353 NIYSPLNNGETLYLFMOQKQSLIFASDINQAGCELYFGCNTTYSRNSQ--TNGQGHYV 411
 QY 418 GERSYEWQVHNPEGRDLSTKIGLTLVNGKGNLSLSVGLVIVLDQADSSGQKQAF 477
 DB 418 GERSYEWQVHNPEGRDLSTKIGLTLVNGKGNLSLSVGLVIVLDQADSSGQKQAF 477
 QY 472 SEIGVSGRGYVQLNDKQPDYDKYFQFGRGR-LDNGSLTFKKIQTDTGAMIVNHT 531
 DB 472 SEIGVSGRGYVQLNDKQPDYDKYFQFGRGR-LDNGSLTFKKIQTDTGAMIVNHT 531
 QY 532 TQAAVYTTIGNESIVLPNGNNINLKLDPYKCEIAYVNGPGETDKNNKRNRLMLIYKPTEDR 591
 DB 532 TQAAVYTTIGNESIVLPNGNNINLKLDPYKCEIAYVNGPGETDKNNKRNRLMLIYKPTEDR 591
 QY 594 HLLSGGTINLNGNITONGGTLVSGRFPYAHVHLRLDLSNNEGIPGELVMDHMDIRI 653
 DB 594 HLLSGGTINLNGNITONGGTLVSGRFPYAHVHLRLDLSNNEGIPGELVMDHMDIRI 653
 QY 652 FKAENFQIKGSSAVSRNYSIEGHWYTSNNAALITGVVNOQNTICTRSDMTGLTTQK 711
 DB 652 FKAENFQIKGSSAVSRNYSIEGHWYTSNNAALITGVVNOQNTICTRSDMTGLTTQK 711
 QY 712 VDLDTKYNSIPKTIQNGSIVLDNATAVNGKGLATNGNTLNLNHSOFTLSNNTQNG 771
 DB 712 VDLDTKYNSIPKTIQNGSIVLDNATAVNGKGLATNGNTLNLNHSOFTLSNNTQNG 771
 QY 772 ILSNSTAVVNDANLNGVHLTDSAQFSLKNSHFSHQIOGKGTVTLLENATWMPBDY 831
 DB 772 ILSNSTAVVNDANLNGVHLTDSAQFSLKNSHFSHQIOGKGTVTLLENATWMPBDY 831
 QY 832 TLQNTLNNSTITLNSAASNNTP--RRSLFTTPTTAAHRENTLTVNGKISGGGT 889
 DB 832 TLQNTLNNSTITLNSAASNNTP--RRSLFTTPTTAAHRENTLTVNGKISGGGT 889
 QY 834 TLQNTLNNSTITLNSAASNNTP--RRSLFTTPTTAAHRENTLTVNGKISGGGT 893
 DB 834 TLQNTLNNSTITLNSAASNNTP--RRSLFTTPTTAAHRENTLTVNGKISGGGT 893

QY 890 FQTSLSLFGYKSDKLKISNDAGDYILSVRNTGKEPTELOUTLVESKONQPLSDKLKT 949
 DB 890 FQTSLSLFGYKSDKLKISNDAGDYILSVRNTGKEPTELOUTLVESKONQPLSDKLKT 953
 QY 950 LENDHYDAGALRYKLVNNGEERFLNRPYKEQELHNDLVABAEKTELEAQAQVPTAKTOT 1009
 DB 950 LENDHYDAGALRYKLVNNGEERFLNRPYKEQELHNDLVABAEKTELEAQAQVPTAKTOT 1013
 QY 1010 GEPKYSRRAARAAPETLPDQSLNALAEKQALTAETQSKAKTKYRSKRAV--FSD 1067
 DB 1010 GEPKYSRRAARAAPETLPDQSLNALAEKQALTAETQSKAKTKYRSKRAV--FSD 1067
 QY 1068 PLDQSLFALAEVYDAAPQSEKDLAQAQ--QKQKDLISRYNSNLSLSEATVN 1126
 DB 1068 PLDQSLFALAEVYDAAPQSEKDLAQAQ--QKQKDLISRYNSNLSLSEATVN 1124
 QY 1127 SMLSYDELDRLFPVDAQSAVWNTNACQKRYNSDAFRAVOQKTNLROTGVOKALANR 1186
 DB 1127 SMLSYDELDRLFPVDAQSAVWNTNACQKRYNSDAFRAVOQKTNLROTGVOKALANR 1183
 QY 1187 IGAVSHSRSDMTFEDQVQKHAATLTWNSGFAQYQMGDLQFGVNVGTGSLASKAEEQSRK 1246
 DB 1187 IGAVSHSRSDMTFEDQVQKHAATLTWNSGFAQYQMGDLQFGVNVGTGSLASKAEEQSRK 1243
 QY 1247 IIRKAINYGNASVQRLQGLGIQYFGVNRFYERENTYQSEEVRYVTPSLAFNRYNAGI 1306
 DB 1247 IIRKAINYGNASVQRLQGLGIQYFGVNRFYERENTYQSEEVRYVTPSLAFNRYNAGI 1303
 QY 1307 RYDYTFEPTNINISVYKRYFFVNVVNSNANQVTVNLTVYQCPRPYQKQVGAKEILHFE 1366
 DB 1307 RYDYTFEPTNINISVYKRYFFVNVVNSNANQVTVNLTVYQCPRPYQKQVGAKEILHFE 1363
 QY 1367 QLSAFISKSQSQSLGKQNVGKLYGRW 1394
 DB 1367 QLSAFISKSQSQSLGKQNVGKLYGRW 1391

RESULT 10
 us-10-080-505-11
 ; Sequence 11, Application us/10080505
 ; Patent No. 6676948
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene, Joseph W.
 ; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
 ; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
 ; CURRENT APPLICATION NUMBER: US/10/080,505
 ; CURRENT FILING DATE: 2002-02-22
 ; PRIOR APPLICATION NUMBER: US 08/296,791
 ; PRIOR FILING DATE: 1994-10-25
 ; PRIOR APPLICATION NUMBER: US 09/839,996
 ; PRIOR FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patent In version 3.1
 ; SEQ ID NO 11
 ; LENGTH: 1391
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 us-10-080-505-11

Query Match 72.8%; Score 5297.5; DB 4; Length 1391;
 Best Local Similarity 74.7%; Pred. No. 0;
 Matches 1052; Conservative 109; Mismatches 216; Indels 31; Gaps 15;

QY 1 MKTIVRLNLTACISLIGIVSQAAMAGHTYEGIDYQYRDBAENKGFYGAQNIKYNNKQ 60
 DB 1 MKTIVRLNLTACISLIGIVSQAAMAGHTYEGIDYQYRDBAENKGFYGAQNIKYNNK 60
 QY 61 GQLVGSKTKAPMDIFSVSNGVAALVENQYIVSAHVNGYTDVDFGAEGNPDQHRFT 120
 DB 61 GEMIGTMKGVPMPDLSSMVRGYSSTLSEQHLLSVAHNGYDVDFGEGENPDQHRFX 120
 QY 121 YKIVKNNYKKNLHFEEDYHNPRLHFKFTBAPPIVTSNNXGSSYSDRTKYPERVYIG 180
 DB 121 YKIVKNNYKKNLHFEEDYHNPRLHFKFTBAPPIVTSNNXGSSYSDRTKYPERVYIG 180

```

121 YKVVXRYXKSGDHQ--YNDYQHPRLEKFTETAPLEBVSYDNGHNYNNQPLAVRNG 178
181 SGRQFMRNDOK--GQVWASHYLLAINTHNGAGNGSYLGGDYRKAKEYGPIPLAG 238
179 SHQWAKODNNKTTIGDIAYG--SWLIGNTEDGDPAGNQTLEINRYQNNKYGPIPLAG 237
239 SKGDSGSMFYDAEKQKWLINGIIREGNPEKENGPOLVKYSVFIERDLHSLYT 298
238 SFQDSGSMFYDEKVKWLVNGVIREGNPAAVNGSYQITKDYFGQILNODITANFMD 297
299 RAGNGVTTISGNDK---QGSITQKSGIPSEIK-TTLANMSL-PLAKDKVANPRYDP 352
293 TWAEYREYIGSDHNGRVAETKSTLPKALQPEALVGLYDNGSLDADADKGDSSPEYKGP 357
353 KTSPLRNLNGSLYFMDQKQSLIFASDINQAGGLYFGNGFTYSNNSQ-TWQAGAHV 411
358 NFWSPALHNGKSYTFDQGTGLTLENNINGAGGLYFGNGFTYSNNSQ-TWQAGAHV 417
412 SENSITVAKVNGVEHRLSKIGKTLVQVQAGNGKSIYVNGKVTLEQADQNGKQAF 471
418 GRESTVEMQVNEPGRISKIGLGLVYNGKGLASSYVNGSLVLDQADSSQKQAF 477
472 SEIGVSGRGTVQLNDKQFDTDKYFGFGGRLDNGHSLTFKRIQNTDEGAMIVNNT 531
478 KGVGVSGRAVQJNADQVDPNNIYFEGRGRLDNGHSLTFKRIQNTDEGAMIVNNA 537
532 TQAKVCTTGESYVLPNNNNINKLDYKELANGRGRTJAKNGKRLLYKPTTDR 591
538 SQTANITTTGNAT---NSCKQJLNKQIDANGFEGKQAKTNRILVWQVPMANEN 593
592 TLISGNTLKGDIOTQKGLFEGSGPTPHAYNHLKRESEMEGIPGELVMDHMIART 651
594 HLLSGNTLNGNTTNGTIVSGRTPPHAYNHLKRLSNMGEIPGELVMDHMIART 653
652 EKAENFQIKGSAVNSRNTSIEGNWYVSNNAATFVVPQNTICTRSDWTGLTQXK 711
654 FKAENFQIKGSAVNSRNTSIEGNWYVSNNAATFVVPQNTICTRSDWTGLTQXK 713
712 VLTDTKYNSIPKTOINGSLTDMNCAVVKGLACNGNTLNHSGPTLSNNTQJGN 771
714 VLTDTKYNSIPKTOINGSLTDMNCAVVKGLACNGNTLNHSGPTLSNNTQJGN 773
772 TSLDNSTATVDMNANGVHLTDSAQFSLKNSHFSHQLODCKTGTTLNATWTMPBDT 831
774 TSLVHANAATVDMNANGVHLTDSAQFSLKNSHFSHQLODCKTGTTLNATWTMPBDT 833
832 TLQNTLTANSTTILANSASASNNTP--RRSLTETTTPTSAERFNTLTVNGKLSGGGT 889
834 TLQNTLTANSTTILANSASASNNTP--RRSLTETTTPTSAERFNTLTVNGKLSGGGT 893
893 FQFTSSLEFGYKSDKLSNDAREGYTILSVNNGKPELTLEGLTVNKSNOPLSKKKT 949
894 FQFTSSLEFGYKSDKLSNDAREGYTILSVNNGKPELTLEGLTVNKSNOPLSKKKT 953
953 TENDVDAGALRYKLVKQDGEFRLNPIKEOELHNDLVRARQAEKTEAKQVEPTAKQT 1009
954 TENDVDAGALRYKLVKQDGEFRLNPIKEOELHNDLVRARQAEKTEAKQVEPTAKQT 1013
1012 CEPKYSRRAARAAPDTLPOOSLNLAKQAEFTATQOSKAKTKYVSRRAV--ESD 1067
1014 SKAVRSR--RAVFSDDLPAOSLNLAKQAEFTATQOSKAKTKYVSRRAV--ESD 1067
1068 PLDQSLFALEALVETIDAPOOSEKDLAQEZAEK--GRKQDLSRYSNLSALSELSTVN 1126
1068 TLPPQ--ILQALVETIDAPOOSEKDLAQEZAEK--GRKQDLSRYSNLSALSELSTVN 1124
1127 SMLSVDLRLFEVDQASAVNTIADQRRYSDAFRAVQOQKTNARQIGVQALANGR 1186
1125 SMLSVDLRLFEVDQASAVNTIADQRRYSDAFRAV--QOQKTNARQIGVQALANGR 1183
1187 IGAVFSHSDNTPEQYKHAATLTWVSGFAQYQMGDLOGVNNGVIGISAKMAEBSRK 1246

```

```

1184 IGAVFSHSDNTPEQYKHAATLTWVSGFAQYQMGDLOGVNNGVIGISAKMAEBSRK 1243
1247 IHRKALVGNASVQFRLGQIGIOPFGVNRFFIERENYOSSEVRYKTESLAFNRVYAGI 1306
1244 IHRKALVGNASVQFRLGQIGIOPFGVNRFFIERENYOSSEVRYKTESLAFNRVYAGI 1303
1307 RVDTTFPDDNISVAPYFFVYVDSNANVOTTNLVLQOPFGRYQKVEGLAAEILHF 1366
1304 RVDTTFPDDNISVAPYFFVYVDSNANVOTTNLVLQOPFGRYQKVEGLAAEILHF 1363
1367 QISAFISKQSGSLGKQOQNGVVLGYRW 1394
1364 QISAFISKQSGSLGKQOQNGVVLGYRW 1391

RESULT 11
US-08-296-791-3
Sequence 3, Application US/08296791
Patent No. 624537
GENERAL INFORMATION:
APPLICANT: St. Geme III, Joseph W.
APPLICANT: Falkow, Stanley
TITLE OF INVENTION: Haemophilus Adherence and Penetration
TITLE OF INVENTOR: Protein
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: Flehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/296, 791
FILING DATE: 25-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Treacart, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59941/REF/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1541 amino acids
TYPE: amino acid
TOPOLOGY: unknown
US-08-296-791-3

Query Match 23.0%; Score 1673.5; DB 3; Length 1541;
Best Local Similarity 28.5%; Pred. No. 1,46-101;
Matches 480; Conservative 246; Mismatches 512; Indels 445; Gaps 56;

6 FRLNFTACISIGVSOAMAGHTYFGIDYQYRDPANNGKGTVCAGKIKYNNQOGLG 65
6 FRLNFTALVAVALTPTTBAALVRODVYQIFRFPANNGKTSVATIVLVKDNKNNDIG 65
6 TSMTKA-EMIDSVYS-RNGVVALVENQIVSYAH-NVGITDVDEG-AEGRN---NPDOH 117
6 TALPAGIMIDFVSVDVDRKIALILNPQYVGVYHVSNGVSLHGNLNGMNNAGAKXH 125
118 RFL-----YKIVKANY-----KQNLHPYEDYIHPRLHFTVTEAPID-MTS 160
126 KVVSEENRYSVSKVEYFTKLNKGTVTTEBQYQKREDDYVWPRLDKVFTEVAPISASTA 185
161 NNKSTYSDRTKTPERVIRIGSGRQFMRNDOKDQ-----VAGAYHVL 203

```


SEQUENCE DESCRIPTION: SEQ ID NO: 3
US-09-839-936-3

Query Match 23.3%; Score 1673.5; D3 4; Length 1541;
Best Local Similarity 28.5%; Pred. No. 1,4e-101;
Matches 480; Conservative 246; Mismatches 512; Indels 445; Gaps 56;

```

QY 6 FPLNLTACISLGVSOAMAGHTFYGDYDYPENKKGKFTVGAQNIKYVNGQGLY 65
D 6 FPLNLTACISLGVSOAMAGHTFYGDYDYPENKKGKFTVGAQNIKYVNGQGLY 65
QY 66 TSMYKA-PMIDSEVYS-RNGVAAVENVQVSYAH-NVGYTDVDFG-ABGN---NPDQ 117
D 66 TSMYKA-PMIDSEVYS-RNGVAAVENVQVSYAH-NVGYTDVDFG-ABGN---NPDQ 117
QY 118 FET-----YKIVKANNY-----KXONIHPYEDDYNNPILKFTYEAAPD-MTS 166
D 118 FET-----YKIVKANNY-----KXONIHPYEDDYNNPILKFTYEAAPD-MTS 166
QY 126 RDVSESENRKFSVEKKEPPTKMGKVTTEEDQTKRRREDYVPRDKFTVEVAPLESTA 185
D 126 RDVSESENRKFSVEKKEPPTKMGKVTTEEDQTKRRREDYVPRDKFTVEVAPLESTA 185
QY 161 NMNGSYSDRTKXPERVAISGRQFMHNDQKQD-----VAGAYHYL 203
D 161 NMNGSYSDRTKXPERVAISGRQFMHNDQKQD-----VAGAYHYL 203
QY 186 SSDAGTYNDQNKYPAFAVLGSGSQF---LYKGDVYSILNNEHVGANNLKLVDGATTYG 242
D 186 SSDAGTYNDQNKYPAFAVLGSGSQF---LYKGDVYSILNNEHVGANNLKLVDGATTYG 242
QY 204 TACVTENQKAGKNGSYTG-----GVYKAGEYGPPLPIAGSKDGSMPLYDAEYQK 257
D 204 TACVTENQKAGKNGSYTG-----GVYKAGEYGPPLPIAGSKDGSMPLYDAEYQK 257
QY 243 IAGPYKVNHNNGGIGFNGSKHSPKGLISODELTNYAVLGDSGSLFYDREKGM 302
D 243 IAGPYKVNHNNGGIGFNGSKHSPKGLISODELTNYAVLGDSGSLFYDREKGM 302
QY 258 LINGILREGNPEEKE-NGFQYVKSYPDEIFERDLHSL-----YTAKNG-VYITSG 309
D 258 LINGILREGNPEEKE-NGFQYVKSYPDEIFERDLHSL-----YTAKNG-VYITSG 309
QY 303 LFQSDYDFMAGNYKKSQWEMVITYSPTKDVANKQASGLSKTDYSSNGKSTYTG 362
D 303 LFQSDYDFMAGNYKKSQWEMVITYSPTKDVANKQASGLSKTDYSSNGKSTYTG 362
QY 310 NDNGQGSITQKSGIPSEIKITLAMSILPKKKQKVNPRYDGNVYSPILNGETLYFM 369
D 310 NDNGQGSITQKSGIPSEIKITLAMSILPKKKQKVNPRYDGNVYSPILNGETLYFM 369
QY 363 GE-----KSLNDVLAD-----GDKR-----NHGSVTF-- 386
D 363 GE-----KSLNDVLAD-----GDKR-----NHGSVTF-- 386
QY 376 QKQGSILFASDINOAGGLYFEGNFTVSPNS-NOTWQAGIHVSENSYTKVNGVEMH 428
D 376 QKQGSILFASDINOAGGLYFEGNFTVSPNS-NOTWQAGIHVSENSYTKVNGVEMH 428
QY 387 EGSGTLNNNIDQAGGLFPEGDEYKGTSDNTWKAGAVYABEKTYWKVNPQYR 446
D 387 EGSGTLNNNIDQAGGLFPEGDEYKGTSDNTWKAGAVYABEKTYWKVNPQYR 446
QY 429 LSKIKGTLHQAENGKSGISYDQKVLLEQADDOQNAQSHGLVSRGVQVOLD 486
D 429 LSKIKGTLHQAENGKSGISYDQKVLLEQADDOQNAQSHGLVSRGVQVOLD 486
QY 447 IAKIGKGTILVEGTQDNKSLKVGDDGYVLKQQTNGSG-QHAFASGVISRSRTLYLND 503
D 447 IAKIGKGTILVEGTQDNKSLKVGDDGYVLKQQTNGSG-QHAFASGVISRSRTLYLND 503
QY 489 KQPDYKXYFEGFRGLDNLNLSLTPKRLQNTDQAMIVNENYTOQAVNTYNGS-VLP 548
D 489 KQPDYKXYFEGFRGLDNLNLSLTPKRLQNTDQAMIVNENYTOQAVNTYNGS-VLP 548
QY 506 KQYDPSIYFEGFRGLDNLNLSLTPKRLQNTDQAMIVNENYTOQAVNTYNGS-VLP 565
D 506 KQYDPSIYFEGFRGLDNLNLSLTPKRLQNTDQAMIVNENYTOQAVNTYNGS-VLP 565
QY 549 N----- 549
D 549 N----- 549
QY 566 NITTPYINIDAPEDNPPAFRIIDQGLYINLENYTYALRKQASTRELPKNGSENEK 625
D 566 NITTPYINIDAPEDNPPAFRIIDQGLYINLENYTYALRKQASTRELPKNGSENEK 625
QY 590 -----GNNIKLDYRKELIYNGMGFETDKKNGRLNIYPTTERKTLISG 597
D 590 -----GNNIKLDYRKELIYNGMGFETDKKNGRLNIYPTTERKTLISG 597
QY 626 WLYMGKTSDEAKVNYNHINBRNGFNGYCE-EGKKNNGNLNVTFKGSEONRELLTG 684
D 626 WLYMGKTSDEAKVNYNHINBRNGFNGYCE-EGKKNNGNLNVTFKGSEONRELLTG 684
QY 598 GYVKGKIDITQYKALFFSGSPYTHAYNHLNRWSENGI-----PGSEIVMDHM 647
D 598 GYVKGKIDITQYKALFFSGSPYTHAYNHLNRWSENGI-----PGSEIVMDHM 647
QY 683 GYVKGKIDITQYKALFFSGSPYTHAYNHLNRWSENGI-----PGSEIVMDHM 736
D 683 GYVKGKIDITQYKALFFSGSPYTHAYNHLNRWSENGI-----PGSEIVMDHM 736
QY 648 INETFAENFOIKGSGAVYS-RWSSISGNTVYNNANNTFGVYENQONTITCRSPWGL 706
D 648 INETFAENFOIKGSGAVYS-RWSSISGNTVYNNANNTFGVYENQONTITCRSPWGL 706
QY 737 INENFALTMVNTSNASLVSGRVNNTISNTLSNKAQVHIY--KTSCTVYCVRSYTYG 794
D 737 INENFALTMVNTSNASLVSGRVNNTISNTLSNKAQVHIY--KTSCTVYCVRSYTYG 794
QY 707 TCGQKVDLTPKVINSPKTOINGSINLTENATANYGLAKNGVNTLTHSQPTLSNNA 766
D 707 TCGQKVDLTPKVINSPKTOINGSINLTENATANYGLAKNGVNTLTHSQPTLSNNA 766
QY 795 VTCITDKLSD-KALNSPNTNLRGNVLTRESANF-VLGANLPGITQSRGNSQ-- 845
D 795 VTCITDKLSD-KALNSPNTNLRGNVLTRESANF-VLGANLPGITQSRGNSQ-- 845
QY 767 TQIGNRLSDNSTATVDNANLNGVYLTSAQSLNNSHSHQIQSGKGTVTLNKAATW 826
D 767 TQIGNRLSDNSTATVDNANLNGVYLTSAQSLNNSHSHQIQSGKGTVTLNKAATW 826
QY 846 -----VRLTENS-----HMLTGN--SDVHQDLNNGH----- 871
D 846 -----VRLTENS-----HMLTGN--SDVHQDLNNGH----- 871
QY 827 MSQDTTLQNLINNSTITLNSAVSSNNTPRRSLETETTTLSAHRNVTLTWNGKLSG 886
D 827 MSQDTTLQNLINNSTITLNSAVSSNNTPRRSLETETTTLSAHRNVTLTWNGKLSG 886

```

```

D 872 -----IHLNSA--DNSNNVT-----KXNLTLYN-SLSC 896
QY 887 QGFQFTSLFPGYSKDKLKLSNDABEDYLLSVRTGKPELTLEQLVSKONQPLSDXL 946
D 887 QGFQFTSLFPGYSKDKLKLSNDABEDYLLSVRTGKPELTLEQLVSKONQPLSDXL 946
QY 897 NGSFYTLTDLMSKQGDVYVTKSANTNFTLQVAKTGPENH-NEULTDASKAQ--RDH 953
D 897 NGSFYTLTDLMSKQGDVYVTKSANTNFTLQVAKTGPENH-NEULTDASKAQ--RDH 953
QY 947 KETLENDYDAGALRYLVANDGEFRLLNPLKEQELHNDLV----- 987
D 947 KETLENDYDAGALRYLVANDGEFRLLNPLKEQELHNDLV----- 987
QY 988 -----BARCARLEAKQVET----- 1004
D 988 -----BARCARLEAKQVET----- 1004
QY 1012 SNNEIARVDAFPAPPAPATSETTETTVANSKQSKTYENKQDATTETTAQREYAK 1071
D 1012 SNNEIARVDAFPAPPAPATSETTETTVANSKQSKTYENKQDATTETTAQREYAK 1071
QY 1005 -----AKTQGE-----PKYSSRAA 1020
D 1005 -----AKTQGE-----PKYSSRAA 1020
QY 1072 AKSNVANTQINEVAQSGSETKBTQTTBTKETATVEKEKAKVETKQVAPKATQSVSP 1131
D 1072 AKSNVANTQINEVAQSGSETKBTQTTBTKETATVEKEKAKVETKQVAPKATQSVSP 1131
QY 1021 RAAPDTLPPDS-----LNLAEAK-QAELTATQSKAKTKKRSKRAVSPDL 1071
D 1021 RAAPDTLPPDS-----LNLAEAK-QAELTATQSKAKTKKRSKRAVSPDL 1071
QY 1132 KQSGELVQPAEPARENDPTVNIKEPOSQNTTADTEOPAKETS-----SNVEQPYTE 1185
D 1132 KQSGELVQPAEPARENDPTVNIKEPOSQNTTADTEOPAKETS-----SNVEQPYTE 1185
QY 1072 QSLFALAEALVIDAPQO-----SEKDRLOEAKQKQK----- 1107
D 1072 QSLFALAEALVIDAPQO-----SEKDRLOEAKQKQK----- 1107
QY 1186 ST--TVNNGSVNENPENTTPATQPTVNSSESKRKRNRHSVRSYFHNVEPATYSSND 1243
D 1186 ST--TVNNGSVNENPENTTPATQPTVNSSESKRKRNRHSVRSYFHNVEPATYSSND 1243
QY 1108 -----DLISRYNSALSLSA-----TVNSMLSYODELDLFLV-DOAQSAVMTNIAD 1154
D 1108 -----DLISRYNSALSLSA-----TVNSMLSYODELDLFLV-DOAQSAVMTNIAD 1154
QY 1244 RSTVALCULSTNTNAVLSDAKAKQFVALNVGKAVASQHIQLSENNENGGVNVNVENTEM 1303
D 1244 RSTVALCULSTNTNAVLSDAKAKQFVALNVGKAVASQHIQLSENNENGGVNVNVENTEM 1303
QY 1155 KERYSDAFRAYQOQNTNLRQIGVQKALANG-RIGAVFESHSDNTDEQVKNHATLYM 1213
D 1155 KERYSDAFRAYQOQNTNLRQIGVQKALANG-RIGAVFESHSDNTDEQVKNHATLYM 1213
QY 1304 NKQISSQYRFSKSTQ--TQGLMDQITISNVVGLGVFTYVNSNNEPKATSKR-TLAQV 1361
D 1304 NKQISSQYRFSKSTQ--TQGLMDQITISNVVGLGVFTYVNSNNEPKATSKR-TLAQV 1361
QY 1214 SGFAQY-----QMGDLOFVNVGTGISASKMAEBSKRIHKAIVNGVNAVYOPRLQGLI 1269
D 1214 SGFAQY-----QMGDLOFVNVGTGISASKMAEBSKRIHKAIVNGVNAVYOPRLQGLI 1269
QY 1362 NYSKYVADNMH---VGLDLSYKGFQSKQITNNAKKEPHTAQGLIAGAKANLQNFPI 1418
D 1362 NYSKYVADNMH---VGLDLSYKGFQSKQITNNAKKEPHTAQGLIAGAKANLQNFPI 1418
QY 1270 QYFVFNRYETTERBNYQSEEVKVTPTSLAFNRVYAGIRVDYTFPTPTNISVKEFFVNY 1329
D 1270 QYFVFNRYETTERBNYQSEEVKVTPTSLAFNRVYAGIRVDYTFPTPTNISVKEFFVNY 1329
QY 1419 TVIVGRVSYLSNADBALDQARIKVNPLSVKTAFAQVDSLTV-HLGEFVPTLSAR- 1476
D 1419 TVIVGRVSYLSNADBALDQARIKVNPLSVKTAFAQVDSLTV-HLGEFVPTLSAR- 1476
QY 1330 DYSNANVQTYMLTVLQCPKRYQKVEGLKAEILHFOISAFISKSGSQGLAQONVGVK 1389
D 1330 DYSNANVQTYMLTVLQCPKRYQKVEGLKAEILHFOISAFISKSGSQGLAQONVGVK 1389
QY 1477 DNNQSGGIVNNGYDPAYNVENQOYNAGLKLKYNHVKSLIGGLYAKQAEKQYAEIK 1536
D 1477 DNNQSGGIVNNGYDPAYNVENQOYNAGLKLKYNHVKSLIGGLYAKQAEKQYAEIK 1536
QY 1390 LGY 1392
D 1390 LGY 1392
QY 1537 ISF 1539
D 1537 ISF 1539

```

RESULT 13

```

US-10-080-505-3
; Sequence 3, Application US/10080505
; Patent No. 6676948
; GENERAL INFORMATION:
; APPLICANT: St. Gene, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
; CURRENT APPLICATION NUMBER: US/10/080,505
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent version 3.1
; SEQ ID NO 3
; LENGTH: 1541
; TYPE: prt
; ORGANISM: Haemophilus influenzae
US-10-080-505-3

```


Query Match 23.0%: Score 1673.5; DB 4; Length 1541;
 Best Local Similarity 28.5%: Pred. No. 1.4e-101;
 Matches 480; Conservative 246; Mismatches 512; Indels 445; Gaps 56;

6 FRNLPLTACISLIGVSOAMAGHTYFGIDYQYRDPFAENKQFTYGAOMIKYVNRQQLV 65
 6 PKCNFLALIVAAALPYTBALVRDQVQIPRPAENKQFTYGAOMIKYVNRQQLV 65
 6 TSMTXA-PMIDSTVS-RGVAAVVENQYIVSVA-NVGTDDVDRG-AEGR---NEDQH 117
 66 TALPNCIEMIDSVYDQKRIATLIMPQYVGVNNGVSEHFGNNGMNGAKAH 125
 118 RFT-----YKIVKNNY-----KCONLHPEYEDVHNPRLKETEVAIID-MNS 163
 126 RDVSSSENNYFVEKQEPYKLNKGTVTTEQOTQKREDDYMERLQKTEVVAIEKSTA 185
 161 NANGSTYSDRTYPERVYISGRQFMRNDQKQD-----VAGAYHYL 203
 186 SSDGTYNDQNKYPAFRLSGSQF---LYKKGNYSLLIHHVEGNNKLVDADATYTG 242
 204 TAGNTNORAGAGNYSTLS-----GVYKAGYGPPLIAGSKGSSPMFYDAEKQY 257
 243 IAGTPYKVHNNHGLIGFNSKEHSDPKILSODPLTNVAVLGSGSGPLFYVDRKQK 302
 258 LINGILREGNPEKGE-NQSQVYKSYFBIFFEDLHSL-----YTRAGNG-VYTI 309
 303 LFTGSDYFAGYNNKSKQWENITYSQFTYDVLNKSAGSLGSKTDYSSNGKSTSTTG 362
 310 NDNGQSLYQSGI-PSEIKITLANNSLPLKEKDYVHNPYDGNITYSRLNGETLYFM 369
 363 GE-----KSLANDLAD-----GKDF-----NHGKSTYF-- 386
 370 QKQSGIIFASDINGAGGLYFEGNFTVSPNS-NOTWQAGIHVSNSTVTWKVNGVHNR 428
 387 EGGGTLNNNNIDQAGGLFFEGDEYVGTBDNTTNGAGVYAAKVTWKVHNQYDR 446
 429 LSKIGKGLHYQAKGENKGISYGGQVILEQADQKQKQFSEIGYVSGRYQYQND 488
 447 LAKISGGLIYEGTGDNKSGLKVGDETVILKQQTNGSG-OHAFASVGVISRSRLTVAND 505
 489 KQFDIDKPYFEGRGRLDNLNGHSLFKRIQNTDEGAMIVHNTTQANAVITIGNESIVLP 548
 506 KQVDPMSIYFEGRGRLDNLNGHSLFKRIQNTDEGAMIVHNTTQANAVITIGNESIVLP 565
 549 N----- 549
 566 NITTPYNIADPEDNPYAFRRIKDGQQLYLNLENTYVALRKAGSTRSELPRKSGSENN 625
 550 -----GNINKLDYRKELIAYKMGAFGETDNKXNGKLNLIYKFTTERILLSG 597
 626 NLYMGKTSDEAKRWVNHINENKNGENGIFGB-BRKNNGNLNVTFRKSGSEONRPLTGG 684
 598 GTNLAKGDIQTQKGLFFSGRPTPHAYNHLNKRSEMEGI-----DQGIIVDHDW 647
 685 GTNLGDIITVEKGLFSGRPTPHAYNHLNKRSEMEGI-----DQGIIVDHDW 736
 648 INRTKRAENFOIKGSGAVVS-RVSSISGAWTYSNNAAKTFVAVPNOQNTITCRSPMTEL 706
 737 INRTKRAENFOIKGSGAVVS-RVSSISGAWTYSNNAAKTFVAVPNOQNTITCRSPMTEL 794
 707 TTCQKVDLDTKVLNS-PRKQINGSLINTDANFANYKGLAKNGVNTLTHNSQFTLSNNA 766
 795 VICTIDKXSD-KALNSRPTNLRGVNLTBSANF-VLGSQANLFGTIQSRNGO----- 845
 767 TQIGIRLSDNSTAYVDNANLNGVNLHLSAQSLKNSHSHQIQDGKGTTLLENATWT 826
 846 -----VRLTENS-----HMLTGN-----SDVHQDLANGH----- 871
 827 MRSDDTLQNTLNNSTITLNSAYSSANNTPRRRSLETFTTTSABERNTLTVNGKLSG 886
 872 -----IHNLSA-----DNSNNVT-----KXVTLTVN-SLSG 896
 887 QGTFOFTSSLFGYKDKLNDAGGYTILSVNRTQKEPTELLQTLTVESKQNPISDKL 946

897 NSSEFYITLDSNNKQSGKVVVTSATNGFTLQVADKXGEENH-NEJLTFDASKAQ--RDLH 953
 947 KFTLNDHVDGALRYKLVKNDGEERLNPFIKQCEHLNLY----- 987
 954 NNSLVGNTVLDGKMKTKLANNNGRIDLYNP-EVEGRQNTVDTTNITTTNNIQTADVSPV 1011
 988 -----RAQOERTLEAKQVEPT----- 1004
 1012 SNNEIARVADPAVPPPADPATPSETTETVVENSKQSKTVEKNEQDATTETAGREVAKE 1071
 1005 -----ATQTCG-----PYRSRRRA 1020
 1072 ANSNKATQOTNEAOSGETYQTTETKETATVEKKAKEVETKTOEPVPTQSVSP 1131
 1021 RAAPDTLPDQS-----LNLALAK-OAELTAETOKSKYAKTKRSRAVAFSPDPLD 1071
 1132 KQSGETIYQPAEPAPENPVTNKEPOSQNTTADTEBPAAKTS-----SVEQPVTE 1185
 1072 QSLFALALAVTDAPQ-----SEKRLAQEBARKQKQ----- 1107
 1186 ST--TVNTGNSVVENDETTPTATQPTVNSSESNKPKDNRHSVSVPHNVEPATSSND 1243
 1108 -----DLRSNSALSELSA-----TNSMLSTQDELDRLPY-DQASAVMTIAD 1154
 1244 KSTVALCDLSTNTNNAVLSDARAQFVALNWKAKASQHSIQLEMMNNGQYNWVNSISM 1303
 1155 KRRYSDAFRAYOQKTNLROIGVQKALANG-RIGAVFHSRSDNTFDEQKHAULTM 1213
 1304 NKNYSOYRFRSSKSTQ-TQLGMDQITISNNVQLGAVFTYRANNNNDKATSKN-TLAV 1361
 1214 SGFAQY-----QMGDLQFGVNVGTGISASRAEBSQKIRKRAINYGNASYOFRIGOLCI 1269
 1362 NFYSXYADNMH---YLGDIGYKGFQSKQTNHNAKAPHTAAGGLTACKAPNAGNGEI 1418
 1270 QPYGVNRYEIEKNNQSEEVKVTPLAPNRKAGIKVDTFPTPNISVYRPNVY 1329
 1419 TVIVGVNRYSTLSNADALDQARIKVNPISVKTAFQVDSITY-HLGEFVTPILSARY- 1476
 1330 DVSNANVQTTNLTVLTQOPFGRYQKEVGEKAEILHPIQSFIFKISQSGSLGKQOQVGVK 1389
 1477 DANGSKIVNVDYDANNVENQOYNAGIKLKYNNKLSLGLYTKAKABEKQKTAELK 1536
 1390 LGY 1392
 1537 LSF 1539

RESULT 14
 PCT-US95-10661A-3
 : Sequence 3, Application PC/TUS9510661A
 : GENERAL INFORMATION:
 : APPLICANT: Washington University, et al.
 : TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
 : NUMBER OF SEQUENCES: 9
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Flehr, Hobbach, Teat, Albritton & Herbert
 : STREET: 4 Embarcadero Center, Suite 340C
 : CITY: San Francisco
 : STATE: California
 : COUNTRY: United States
 : ZIP: 94111-4187
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US95/10661A
 : FILING DATE: 16-AUG-1995
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/296,791

FILING DATE: 25-AUG-1994
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Treacartin, Richard F.
 REGISTRATION NUMBER: 31,801
 REFERENCE/DOCKET NUMBER: FP-59941/RET
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 398-3249
 TELEFAX: (415) 781-1989
 TELETYPE: 910 277239
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1541 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 PCT-US35-10651A-3

Query Match 23.0%; Score 1673.5; DB 5; Length 1541;
 Best Local Similarity 28.5%; Pred. No. 1,4e-101;
 Matches 480; Conservative 246; Mismatches 512; Indels 445; Gaps 56;

QY 6 RLNFELTACISLGVISQVAGHTYEG:DVGYVDRPAENKKGKFTTGAGN:KYVNRQGGQVVG 65
 D 6 EKLNFALTYAVALTYTEALVDDVDVQIFEDPAENKKGKFSVGAATVYVQKNNKDLG 65
 QY 66 TSMTKA-PMIDESVVS-RNGVVALVENQYIVSAH-NVGYTDVDFG-AGEG-N-NDPOH 117
 D 66 TALNGIPMIDFESVDVDRKIALTLINPQYVGVGVNGVSELHFGKLNKNNNNKAKH 125
 QY 118 RFL-----KYVKNNT-----KKNLHPYEDDYENPRLHKFTEAPID-MTS 160
 D 126 RDVSEENRATFSEKKEYPITLNGKVTTEDTQKREDDYWPRLDKFTEVAPLEKSTA 185
 QY 161 NMNSSTYSRKYPERVATISGRQFRNNDGQD-----VAGATAYL 203
 D 186 SSDAGTYNDQKTPAFVLSGSQF---YKKGDNYSILNNEHVGANNKLVGQATYTG 242
 QY 204 TASNENQKAGNGSYLYG-----GPRKAGEVGLPIAGSKDGSMPFIYDAEKOKM 257
 D 243 IAGPYKVNHNKGLIGFNGSKHEHSDPKGLSDQPLNVAVLGDSGSLPYDREKGM 302
 QY 258 LINGILREGNPEEGE-NGFQVYKSYFEIIFERDLTSL-----YTRAGN-VYTTISG 309
 D 303 LPLDSYDFWAGNKKSKWQEMKIYKQFTDVLNKSAGSLGSKTYSSWNGKSTYTG 362
 QY 310 NDNGQGSITQKSGIPSEIKITILANMSLPKEDKYNHPRYDGNITISPLNNGETLYFMD 369
 D 363 GE-----KSLNVDLAD-----GKNP-----NKGKSVTF-- 386
 QY 370 QKQSLIFASDINOGAGLYEENFTVSPNS-NOTWQAGIHVSNSYTWKXNGVEHDE 428
 D 387 EGSGLTLANN:IDQAGGLFPEGDEYVKTSLNLTWKGAGVSAEKKVTKKYNRPQYDR 446
 QY 429 LSKTGGTALHQAQKENGKSI SVGDGKVLLEQADDOQNAQFSEIGLVSRGTVOLND 488
 D 447 LAKTGGTALVEGTGDNKSGSLKVDGIVLTKQNTGSG-OHAFASGIVASRSTVLND 505
 QY 489 KQPDIKFYFGRGRLDINSHSLTEKRLONTDGGAMVNHNTQAQAVTTTGGESLYLP 548
 D 506 KQVDPNSIYFGRGRDLNKGSLTDPDHINIDDAKLVNHNNTNNSITTTGESSLTDP 565
 QY 549 N----- 549
 D 566 NITPPYNDAPDEDNPAFRRIKQGGQLYLNLNMYTYALRKASRSELPJNSGESNEN 625
 QY 550 -----GNINKLDYRKELIYVNGMGEYDKKRNKGLNLIYFTEDETLISG 597
 D 626 KLYNGKTSDEAKRVNVMHNERNGFNQYFG-EEKKNGNLNTVTFKSGEQRFLING 684
 QY 598 GYNLKGDTTOTKGLFSGRPTPHAVNHLNRKMSMGCI-----PGELVMDHW 647
 D 685 GTVNLGDLTVEKGLTFLISGRPTHA-----RDIAGISLSKPKDPHREKNEVEVDDW 736

QY 648 INTRKRNFOIKGSAVVS-RNVSSTEGNTVSNANNAITFVGNQNTICTRSMTGL 706
 D 737 INRFAATMTNTYASISYGRVANITISNITASNKAQVHIG--KTGDTVCVRSDYTG 784
 QY 707 TTCQKVLDPYVINSIPKQVINGS:NLTDNATNVGSLAKANGVTLTNHSTOTLSNA 766
 D 795 VTCTTKKSD-KAANSNPTNLKGVNNTLESANP-VLGKANIFGTIGRSGSQ----- 845
 QY 767 TQIGTRLSDNSTATVNNANNGVNLNDSAQFSLKSHSHQIQGCKATVTLNNAWT 826
 D 846 -----VLTENS-----HMLTGN--SDVHQDLNAGH----- 871
 QY 827 MPEDTTLQNLINNSTITLNSAVSASNNTPRSS:LETTPTSAHRENTLYVNGKLSG 886
 D 872 -----IHLNSA--DNSNVT-----KNYTLTVN-SLIG 896
 QY 887 QGTFOFTSLFGYKSDKLKLSNDABGDYLSVRVYKGPETLEQLTVESKDNPISDLK 946
 D 897 NGSFYTLIDLSKQGDYVVTTSATGNFTLVQADKTGEPKH-NELTFDASKAQ--RDHL 953
 QY 947 KFTLENDVDAALRYKLVDNDEFFLNPIKQELHNDLV----- 987
 D 954 NVSLVGNATVDLGAMKTYLNVNNGRYDLYNP--EVEKNGQVDTNITTPNNIQADVSVP 1011
 QY 988 -----RAGQARTLEAKQVEPT----- 1004
 D 1012 SNNSEIARVDAPVPPAPATSEETETVAENSQKSKVTEREDATETTAQNFVAYE 1071
 QY 1005 -----AKTQGE-----PKYRSRAA 1020
 D 1072 AKSNWANTQVNEAQSGETKETQTTKETATVEKKAVENTEKTQEVPRKTSQVSP 1131
 QY 1021 RAAPFDYLPDS-----LNLAEAK-QAELTAETOKSKATKRVASKRAVSPILD 1071
 D 1132 KQOSSTVQPAEPAENDEPTVNIKEPOSQNTTADTBQPAKETS-----SNVEDPYE 1185
 QY 1072 QSLPALLEALVIDAPQ-----SEKDLAQEAKQKOK----- 1107
 D 1186 ST--TWTGNSVSNPEANTPATTPQTVNSESNNKRNHRSSVSPVHNVPEATTSSND 1243
 QY 1108 -----DLISRYNSALSLSA-----TVNSMLSYODELDLPV-DQASAVWNTIAD 1154
 D 1244 RSTVALCDLSTNTNNAVLSDAKQAPALNVKASQHSISOLENNEGQYVNSISM 1303
 QY 1155 KRRYSDAFRAYQOQKTNLQIGVQKALNG-RIGAVFSHSDNTEDEQYKNAHTLTM 1213
 D 1304 NKRYSSQYRFSKSTQ--TQGMQDTISNNVQLGVFTYVNSNNDKATSKG-TLAQV 1361
 QY 1214 SGFAQY-----QMDLDQGVNVTGTSASRMAEQQSKIKHRAINYGNASYQPRIGQIGI 1269
 D 1362 NFYSKTYADNEM--YLGIDLYGKPOSKLQTHNNAKFPARITAGFLTACKAFNIGFI 1418
 QY 1270 QPFGVNRYPFERENTQSEBEVYKTPSLAFNRVAGIRVDTFPTDNI SVKPYPMVY 1329
 D 1419 PTIVGVRSYLSNADALDQALIKNPIKVPKFAVQVLSITTY-HLGEFVTPILISARY 1476
 QY 1330 DYSANAVOTTYNLTVLQDPGRYQKVEGLKAILIIFQISAFISKSGSLGKQGVGK 1389
 D 1477 DMOQSGKINWGYDPAVVENQOYVNAGLKYHNVKLSLGLTVAKQAEKQVTAELK 1536
 QY 1390 LGY 1392
 D 1537 LSF 1539

RESULT 15
 US-08-296-791-4
 ; Sequence 4, Application US/08296791;
 ; Patent No. 6245317
 ; GENERAL INFORMATION:
 ; APPLICANT: St. Gene III, Joseph W.
 ; APPLICANT: Falow, Stanley
 ; TITLE OF INVENTION: Haemophilus Adherence and Penetration

```

/ TITLE OF INVENTION: Protein
/ NUMBER OF SEQUENCES: 9
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Flehr, Hobach, Test, Albitton & Herbert
/ STREET: 4 Embarcadero Center, Suite 3400
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: United States
/ ZIP: 94111-4187
/ COMPUTER READABLE FORM.
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/296,791
/ FILING DATE: 25-AUG-1994
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Treccartin, Richard F.
/ REGISTRATION NUMBER: 31,801
/ REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
/ TELEPHONE: (415) 781-1989
/ TELEFAX: (415) 398-3249
/ TELEX: 910 277299
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1545 amino acids
/ TYPE: amino acid
/ TOPOLOGY: unknown
/
US-08-296-791-4

Query Match      22.9%; Score 1662.5; DB 3; Length 1545;
Best Local Similarity 28.6%; Pred. No. 7.7e-101;
Matches 462; Conservative 246; Mismatches 511; Indels 447; Gaps 59;

```

```

QY 486 NDDKQFDTKFPGRCGRDLNLSHSLTEPKLQONTDEBAMTVNHTQAANTITGNEI 545
DB 507 NDDKQVDPNSIVYFEGRCGRDLNLSHSLTEPKLQONTDEBAMTVNHTSJKISTITIGDWLI 566
QY 546 VLPNG----- 550
DB 567 TDPNNVSIYVKELEDNPAIROIKYQOLYNEENRTYALKKQASIRSEFPQRGES 626
QY 551 -----NNINELDYRKEIAVNGMPGETDKNENGRILNLYKPTEDRTLL 594
DB 627 NNSLWNGTEKQDAQKAMNHNINERNNGPGRPE-EGGNGNINLVFKKSKQRF 685
QY 535 LSGGTNLKSDITQYTKLPSFGREPTAHNLNKKWSEMEI-----PQGIWMD 644
DB 686 LYGGINMGLDWQGTLELSGRPTPHA-----RDIAGISTKDKSHFSENNVVE 737
QY 645 HDWINTREKRNPOJKGSAAVS-RVSSIEGWNVTYSNANATFPGVVPQQTITCRSD 703
DB 738 DDMWNRFPATNINVTNNATLYSGRVESITSNITASNAKHIGY--YAGDTVCVRSDY 795
QY 704 TGLTTCQKVDLIDTVVINSLEPKTQINGSINTLNDNATAVYGLAKLNGVTLTNHSGFTLS 763
DB 796 TGYVTCITDKSD-KALNSPPTNLGNVNLTESANF-VLGKANLGTITQSKGNSQ--- 849
QY 764 NNATQIGTRLSDNSTATVDNANLNGNHTDLSAQFSLKNSHFSHQIQGDKGTVYLEMA 823
DB 850 -----VALTEMS-----HMLTGN--SDVEQLDLANGH----- 875
QY 824 TWTPSDTTLQWLLNNSITLINSAYSASNNTPRRSLTETTPPSAHRNRTLVNKG 883
DB 876 -----HLNSA--DSSNNVT-----KNTLVN-S 897
QY 884 LSGQGTPTSSLSFGYKSDYLKLSNDAEGDYLSVRNTGKEPETLEQLVEYSDKNOPUS 943
DB 898 LSGNGSFYLLDLSNKGQDVVVYTKAIGNFTLVQAKTCEPHN-NEELTLDASKAQ--R 954
QY 944 DKAKFTLENDVDGALRYLQVKNDEGFRNLNPKIQEQLHNDLY----- 987
DB 955 DILNLSVLTGNTVDLGAKKYLDLRVNGRYDLNVP--EVEKKNQVTDITNITPNNIQADVP 1012
QY 988 -----RAQORLELEKQVEPI----- 1004
DB 1013 SVPSNNEELARDAPPPAPATPSETTEVNAENSKQSKTEVNEBQDAVETTAQNRREV 1072
QY 1005 -----AKTQGE-----PKYRSR 1017
DB 1073 AKKAKSNVANTQINVEAAGSGSETKETQTEKTLATVEKEBAKAVENETKQVEPKTQ 1132
QY 1018 RAARAPPTLPPDS-----LNLALAK-QALETATQSKAKTKYKRSKRAVESP 1068
DB 1133 VSPKQSGETVQPAEPARENDPTVNIKEPOSQNTNTTADTEQPAKETIS-----SNVEOP 1186
QY 1069 LLDQSILALEALEVIDAPQ-----SEKDRLAQEEAEKQKXK----- 1107
DB 1187 VIEST--TYWNGNVANPENPTTPTPTPTNASSSKKRNKRHRSSVTSVHNVEPATTS 1244
QY 1108 -----DLISRYNSALSLSA-----TVNSM.SYODELDLFV-DOAGSAVWTNI 1151
DB 1245 SNDRSTVALCDLITSTYNTAVLSDAKAKQFVALNVGAVASQIHISQLBNNNGQVNVWEN 1304
QY 1152 AADKRRYSDAFRAYQOQNTNLRQGVQKALANG-RIGAVPSHSNDTPEQVKHATL 1210
DB 1305 TSNKNNYSSQYRFRFSKSTQ--TQGMQDQITSNVQGLGVFTYVNSNENPKDKTSKQ--TL 1362
QY 1211 TWMSGFAQY-----QWEDLQGVNVGTGISAKAEQSKIRKXAINVGNASVYQRLQ 1266
DB 1363 AQVNYYSKYVADNHN---YLSDIDLGKFGSKQTMHNKPKRKHADPGLTAKAKANLON 1419
QY 1267 LGIQYPPVNRNYFTEBNTQSEEVVVKTPSLANRNYAGIRVDYTFPTPTNISVKEYPFV 1336
DB 1420 FGITPIVGVRSYSLSNADPALDQARIKVNPLSVKTAFAQVDSLSTY-HLGEFVPTIILSA 1478
QY 1327 NYVDVSANANQTTVNLTLVQPFQGRVYQKXGAKELIHFQISAFISKSGSGLKQQNV 1386

```

Db	1479	RY-DANQSGSKINWNGYDFAYVVENOQYNAIGHKIKYHVKLS.TGGLTAKQAEKQKTA	1537
Qy	1387	GYKLSY	1392
Db	1538	ELKLSF	1543

Search completed: June 23, 2004, 17:30:38
job time : 34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw mode

Run on: June 23, 2004, 17:31:53 ; Search time 63 Seconds
(without alignments)
6559.087 Million cell updates/sec

Title: US-10-645-655-2

Perfect score: 7273

Sequence: 1 MKKVFRLNPLTACISISIV.....SQSGQKQNVGKLGVM 1394

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 282313646 residues

Total number of hits satisfying chosen parameters: 1-63542

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications At: *
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/PTC_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubpaa/PTCUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubpaa/US09D_PUBCOMB.pep:*
13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubpaa/US10D_NEW_PUB.pep:*
17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7273	100.0	1394	10	US-09-839-996-2
2	7273	100.0	1394	12	US-10-645-655-2
3	7273	100.0	1394	14	US-10-083-505-2
4	7246.5	99.6	1395	14	US-10-083-505-7
5	6115	84.1	1434	14	US-10-080-505-9
6	5827.5	83.1	1411	14	US-10-080-505-17
7	5807	79.8	1436	14	US-10-080-505-13
8	5303.5	72.9	1391	14	US-10-080-505-15
9	5297.5	72.8	1391	14	US-10-080-505-11
10	4178	57.4	1449	12	US-10-282-122A-65711
11	4085	56.2	1457	16	US-10-220-481-71
12	4072.5	56.0	1468	12	US-10-282-122A-65027
13	4044.5	55.6	1439	16	US-10-223-481-73
14	1673.5	23.0	1541	10	US-09-839-996-3
15	1673.5	23.0	1541	12	US-10-645-655-3

16	1673.5	23.0	1541	14	US-10-080-505-3	Sequence 3, Appli
17	1662.5	22.9	1545	10	US-09-839-996-4	Sequence 4, Appli
18	1662.5	22.9	1545	12	US-10-645-655-4	Sequence 4, Appli
19	1662.5	22.9	1545	14	US-10-080-505-4	Sequence 4, Appli
20	1615.5	22.2	1593	12	US-10-282-122A-65262	Sequence 65262, A
21	1609	22.1	1694	12	US-10-282-122A-58377	Sequence 58377, A
22	1605	22.1	1702	10	US-09-839-996-5	Sequence 5, Appli
23	1605	22.1	1702	12	US-10-645-655-5	Sequence 5, Appli
24	1605	22.1	1702	14	US-10-080-505-5	Sequence 5, Appli
25	1551	21.3	1848	10	US-09-839-996-6	Sequence 6, Appli
26	1551	21.3	1848	12	US-10-645-655-6	Sequence 6, Appli
27	1551	21.3	1848	14	US-10-080-505-6	Sequence 6, Appli
28	1536.5	21.1	1773	12	US-10-282-122A-65817	Sequence 65817, A
29	1089	15.0	1371	14	US-10-238-075-891	Sequence 891, App
30	1051	14.5	1376	14	US-10-238-075-560	Sequence 560, App
31	346.5	4.8	73	14	US-10-080-505-44	Sequence 44, Appl
32	344.5	4.7	1569	12	US-10-282-122A-43126	Sequence 43126, A
33	344.5	4.7	1569	14	US-10-287-27A-312	Sequence 312, App
34	326.5	4.5	1432	12	US-10-282-122A-78451	Sequence 78451, A
35	292.5	4.0	3755	12	US-10-282-122A-77944	Sequence 77944, A
36	288	4.0	836	12	US-10-282-122A-56613	Sequence 56613, A
37	286.5	3.9	978	12	US-10-092-880-2	Sequence 75356, A
38	285.5	3.9	1536	13	US-10-092-880-2	Sequence 2, Appli
39	285.5	3.9	1004	14	US-10-193-764-53	Sequence 53, Appli
40	285.5	3.9	1010	14	US-10-193-764-51	Sequence 51, Appli
41	283.5	3.9	1222	14	US-10-193-764-37	Sequence 37, Appli
42	283.5	3.9	1228	14	US-10-193-764-34	Sequence 34, Appli
43	283.5	3.9	1536	13	US-10-193-764-63	Sequence 63, Appli
44	283	3.9	1477	14	US-10-092-880-4	Sequence 4, Appli
45	280	3.8	1477	14	US-10-193-764-67	Sequence 67, Appli

ALIGNMENTS

RESULT 1
US-09-839-996-2
Sequence 2, Application US/09839996
Publication No. US2003009010A1
GENERAL INFORMATION:
APPLICANT: St. Gene III, Joseph W.
Falkow, Stanley
TITLE OF INVENTION: Haemophilus Adherence and Penetration
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Fleury, Holbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #2.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/839,996
FILING DATE: 20-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/296,791
FILING DATE: 25-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Treacette, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59941/RT/RTS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:

US 1064565509P1



Creation date: 07-15-2004
Indexing Officer: SMOHAMMED - SUAD MOHAMMED
Team: 1600PrintWorkingFolder
Dossier: 10645655

Legal Date: 07-07-2004

No.	Doccod	Number of pages
1	SRNT	1

Total number of pages: 1

Remarks:

Order of re-scan issued on